AB

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N

A2

(11) International Publication Number: **WO 96/17924**

(43) International Publication Date: 13 June 1996 (13.06.96)

(21) International Application Number: PCT/US95/15463

(22) International Filing Date: 4 December 1995 (04.12.95)

(30) Priority Data:
08/349,498 2 December 1994 (02.12.94) US

(71) Applicants: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US). UNIVERSITY OF WASHINGTON [US/US]; Suite 200, 1107 N.E. 45th Street, Seattle, WA 98105 (US).

(72) Inventors: BEACHY, Philip, A.; 5703 Chilham Road, Baltimore, MD 21209 (US). MOON, Randall, T.; 18531 57th Avenue N.W., Seattle, WA 98155 (US). PORTER, Jeffrey, A.; * (US).

(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

(57) Abstract

The present invention provides two novel polypeptides, referred to as the "N" and "C" fragments of hedgehog, or N-terminal and C-terminal fragments, respectively, which are derived after specific cleavage at a G↓CF site recognized by the autoproteolytic domain in the native protein. Also provided are methods of use of the N and C fragments.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

BACKGROUND OF THE INVENTION

1. *Field of the Invention*

This invention relates generally to the field of protein processing and protein signalling pathways and specifically to two novel proteins having distinct activities, which are derived from a common hedgehog protein precursor.

2. *Description of the Related Art*

Embryologists have long performed experimental manipulations that reveal the striking abilities of certain structures in vertebrate embryos to impose pattern upon surrounding tissues. Speculation on the mechanisms underlying these patterning effects usually centers on the secretion of signaling molecule that elicits an appropriate response from the tissues begin patterned. More recent work aimed at the identification of such signaling molecules implicates secreted proteins encoded by individual members of a small number of gene families. One such family of proteins which may have an influential effect upon patterning activities are those proteins encoded by the *hedgehog* gene family.

The *hedgehog (hh)* gene was initially identified based on its requirement for normal segmental patterning in *Drosophila* (Nüsslein-Volhard, C. & Wieschaus, E, *Nature* 287:795-801, 1980). Its functions include local signaling to coordinate the identities of adjacent cells within early embryonic segments (Hooper, J.E., & Scott, M.P. *Early Embryonic Development of Animals*, pp.1-48, 1992) and a later function in cuticle patterning that extends across many cell diameters (Heemskerk, J. & DiNardo, S., *Cell*, 76:449-460, 1994). The *hh* gene also functions in the patterning of imaginal precursors of adult structures, including the appendages and the eye (Mohler, J. *Genetics*, 120:1061-

1072, 1988; Ma, *et al., Cell*, 75:927-938, 1993; Heberlein, *et al., Cell*, 75:913-926, 1993; Tabata, T. & Kornberg, T.D., *Cell*, 76:89-102, 1992; Basler, K. & Struhl, G., *Nature*, 368:208-214, 1994). Genetic and molecular evidence indicates that *hedgehog* proteins are secreted and function in extracellular signaling (Mohler, J., *supra*; Lee, *et al., Cell*, 71:33-50, 1992; Taylor, *et al., Mech. Dev.*, 42:89-96, 1993).

In vertebrates activities encoded by *hh* homologues have been implicated in anterior/posterior patterning of the limb (Riddle, *et al., Cell*, 75:1401-1416, 1993; Chang, *et al., Development*, 120:3339, 1994), and in dorsal/ventral patterning of the neural tube (Echelard, *et al., Cell*, 75:1417-1430, 1993; Krauss, *et al., Cell*, 75:1431-1444, 1993; Roelink, *et al., Cell*, 76:761-775, 1994).

The vertebrate ventral midbrain contains neurons whose degeneration or abnormal function are linked to a number of diseases, including Parkinson's disease and schizophrenia. It is known that motor neurons develop in close proximity to the floor plate in the ventral midbrain. Midbrain projections to the striatum are involved in the control of voluntary movement (Bjorklund and Lindvall, In: *Handbook of Chemical Neuroanatomy*, eds., Borklund, et al., Amsterdam: Elsevier, pp55-122, 1984) and loss of these neurons results in the motor disorders of Parkinson's disease (Hirsch, *et al., Nature*, 334:345, 1988). Midbrain dopaminergic neurons that innervate limbic structures and the cortex influence emotional and cognitive behavior, respectively, and abnormal function of these neurons has been associated with schizophrenia and drug addiction (Seeman, *et al., Nature*, 365:441, 1993).

While the molecular nature of the factors that specify neuronal cell fate have not been established, members of the transforming growth factor-β (TGF-β) (Lyons, *et al., Trends in Genetics*, 7:408, 1991) or the hedgehog protein family (Smith, J.C., *Cell*, 76:193, 1994) may possess the characteristics expected from such factors as they participate in specification of cell fate, mediate inductive interactions between tissues, and in many cases act at a distance of only a few cell diameters.

The present invention establishes that *hh* activities encoded by these genes play a crucial role in early patterning of the developing eye and in patterning of the brain. For the first time, the invention shows that internal cleavage of hedgehog protein product is critical for full function, and that the two novel products of this auto-proteolytic cleavage display distinguishable activities, thus demonstrating that *hh* signaling activity is a composite effect of two separate signaling proteins that derive from a common *hh* protein precursor. In so doing, the invention provides the means for specific patterning and proliferation of desired neuronal cell types for addressing disorders which arise from neuronal degeneration or abnormal function.

SUMMARY OF THE INVENTION

The present invention is based on the seminal discovery that hedgehog proteins undergo auto-proteolytic cleavage which results in two separate proteins having distinct functional and structural characteristics. The two polypeptides, referred to as the "N" and "C" fragments of hedgehog, or N-terminal and C-terminal fragments, respectively, are produced after specific cleavage at a G$^{\downarrow}$CF site recognized by the autoproteolytic domain in the native protein.

Thus, in one embodiment, the invention provides a substantially pure polypeptide characterized by having an amino acid sequence derived from amino terminal amino acids of a hedgehog protein and having at its carboxy terminus, a G$^{\downarrow}$CF cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide.

In another embodiment, the invention provides a substantially pure polypeptide characterized by having an amino acid sequence derived from carboxy terminal amino acids of a hedgehog protein and having at its amino terminus, a G$^{\downarrow}$CF cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide.

The invention also provides a method for modulating proliferation or differentiation of neuronal cells, comprising contacting the cells with a hedgehog polypeptide. The native hedgehog polypeptide, the N, or the C fragment, or functional fragments derived therefrom, are most useful for the induction of proliferation or differentiation of neuronal cells substantially derived from floor plate neuronal cells.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1 shows processing of the hh protein by immunoblots (A,C) with antibodies against amino (Ab1) and carboxy-terminal (Ab2) epitopes. FIGURE 1B and D are blots of samples immunoprecipitated with Ab1 (B, lanes 7-9), Ab2 (D, lanes 19-21), or pre-immune serum (B, lanes 10-12, and D, lanes 22-24).

FIGURE 1E shows a schematic illustration of the hedgehog cleavage mechanism.

FIGURE 2 shows sequence similarity between hh proteins and serine proteases. hh protein sequences are aligned to residues 323 to 329 of the D. melanogaster protein and numbered as positions 1 to 7 (group A). The catalytic histidines of mammalian serine proteinases (group B) are aligned to the invariant histidine at position 7 in hh proteins.

FIGURE 3 shows autoproteolysis of the hh protein. 3A shows a coomasie blue stained polyacrylamide gel showing production and purification of His_6-U and His_6-U_{H329A} proteins from *E. coli*. Samples were molecular weight markers (lanes 1 and 2); lysates of *E. coli* cells carrying the His_6-U expression construct without (lane 3) and with (lane 4) induction by IPTG; purified His_6-U protein (lane 5); lysates of *E. coli* cells that carry the His_6-U_{H329A} expression construct without (lane 6) and with (lane 7) induction by IPTG; purified His_6-U_{H329A} protein (lane 8). FIGURE 3(B) is an immunoblot detected with Ab2 showing transfected S2 cells induced to express *hh* (lane 1); His_6-U and His_6-U_{H329A} proteins incubated in cleavage reaction buffer for 0 hours (lanes 2 and 5), for 20 hours (lanes 3 and 6), and for 20 hours in the presence of 20 mM TAME (a serine protease inhibitor) (lanes 4 and 7).

FIGURE 4 shows autoproteolytic functions of *Drosophila* (4A-C) and zebrafish (D) hh proteins map to the carboxy terminal fragments by in vitro translations of wild-type and mutant hh proteins. The locations of mutations and cleavage sites (arrows) in these proteins are illustrated schematically in 4E.

FIGURE 5 shows immunoblots showing heat shock induced expression of wild type and H329A mutant hh proteins in *Drosophilia* embryos (A) and (B) are immunoblots developed using Ab1 and Ab2 antibodies, respectively. Lanes 1 and 6, induced untransfected S2 cells; lanes 2 and 7, transfected S2 cells induced to express *hh*; lanes 3 and 8, heat shocked wild-type embryos; lanes 4 and 9, heat shocked hs*hh* embryos; lanes 5 and 10, heat shocked hs*hh* H329A embryos.

FIGURE 6 shows *in situ* hybridization showing the embryonic effects of ubiquitously expressed wild type and H329A hh proteins. FIGURE 6 shows the embryonic distribution of *wingless* (*wg*) RNA as revealed by in situ hybridization is shown in (A) wild-type (homozygous y^1 w^{1118}), (B) hs*hh*, and (C) hs*hh* H329A embryos that were exposed to two 10 minute heat shocks separated by a 90-minute recovery period (33). Wild-type embryos showed little change in *wg* expression, whereas the wild-type protein and, to a lesser extent, the H329A protein each induced ectopic *wg* expression (Table 1). Panels (D), (E), and (F) show the dorsal surfaces of y^1 w^{1118}, hs*hh*, and hs*hh* H329A larvae, respectively, at the level of the fourth abdominal segment. These larvae were shocked for 30 minutes as embryos and allowed to complete embryogenesis. Cuticle cell types (1°, 2°, 3°, and 4°) are labeled as described (J. Heemskerk and S. DiNardo, *Cell* 76, 449, 1994). Note the expansion of 2° cell types (naked cuticle) at the expense of 3° and some 4° types in the hs*hh* embryo (E) under conditions where the phenotype of hs*hh* H329A embryos (F) is identical to that of control embryos (D).

FIGURE 7 shows X-gal staining to show imaginal disc effects of ubiquitous wild type and H329 hh proteins. X-gal staining was used to follow expression of *wg* (A-C) or *dpp* (D-L) in imaginal discs of late third-instar larvae that carry *wg-lacZ* or *dpp-lacZ* reporter genes. Leg (A-F), wing (G-I) and eye-antennal discs (J-L) from control larvae (A, D, G, J), larvae carrying the hs*hh* transgene (B, E, H, K) and larvae carrying the hs*hh* H329A transgene (C, F, I, L) are displayed. In all panels anterior is to the left.

FIGURE 8 (A) and (B) are immunoblots of cell pellets (lane 1) or supernatants (lane 2) from transfected S2 cell cultures expressing HH protein, developed with Ab1 (A) and Ab2 (B). Samples in each lane were from the same volume of resuspended total culture. Whereas N remained mostly associated with the cell pellet (compare lanes 1 and 2 in A), C was nearly quantitatively released into the supernatant (compare lanes 1 and 2 in B). U displayed partitioning properties in between those of N and C (A and B). (C) demonstrates the heparin binding activity of various HH protein species generated by in vitro translations with microsomes (38). Samples were: total translation mix (lane 1); supernatant after incubation with heparin agarose or agarose (control) beads (lanes 2 and 4); and material eluted from heparin agarose or agarose beads after washing (lanes 3 and 5). F, U, N_{ss} and N fragments are depleted from reactions incubated with heparin agarose but not agarose beads (compare lanes 2 and 4 to 1), and the same species subsequently can be eluted from the heparin agarose but not the agarose beads (compare lanes 3 and 5 with lane 1).

FIGURE 9 shows the differential localizations of N and C in embryos by in situ localization of the *hh* transcript. Fig. 9 (A) is shown in comparison to the distribution of N and C epitopes detected with Ab1 and Ab2 in panels (B) and (C), respectively. Note that the distribution of N and C epitopes span approximately one-third and one-half of each segmental unit respectively, while the transcript is limited to approximately one-quarter of each unit. In (D), the localization of C epitopes in embryos homozygous for the hh^{13E} allele is detected with the use of Ab2. C epitopes in this mutant, which displays impaired auto-proteolytic activity (see text), are more restricted, and resemble the wild-type localization of N. Homozygous hh^{13E} embryos were identified by loss of a marked balancer from a heterozygous parent stock. All embryos are at mid to late stage 9 (extended germ-band).

FIGURE 10 shows a signal relay versus dual function models for hh protein action. In Fig. 10 (A), the long-range effects of hh signaling are achieved indirectly through short-range induction of a second signaling molecule (X). Based on its biochemical properties

and its restricted tissue localization, N is presumed to represent the active short-range signal while the role of C would be limited to supplying the catalytic machinery required for biogenesis of N. In (B), the long- and short-range signaling functions of hh are supplied by the N and C proteins derived by internal auto-proteolysis of the U precursor. N is implicated in short-range signaling by retention near its cellular site of synthesis, while C is less restricted in its distribution and would execute long-range signaling functions. In both models, auto-proteolysis is required to generate fully active signaling proteins.

FIGURES 10 C and D show an immunoblot of the N fragment synthesized from a wild type construct (C) or a consruct lacking the C domain (D).

FIGURES 11 A and B show the nucleotide and deduced amino acid sequences for partial human *hh* clones.

FIGURE 12 A and B show *in vitro* cleavage reactions of a *Drosophila hh* protein produced in *E. coli* and purified to homogeneity. FIGURE 12, Panel A shows a time course of cleavage after initiation by addition of DTT. Panel B shows incubations of concentrations ranging over three order of magnitude for a fixed time period (four hours), with no difference in the extent of conversion to the cleaved form. Panel C shows the sequence around the cleavage site as determined by amino-terminal sequence of the cleaved fragment C. The cleavage site is denoted by the arrow, and the actual residues sequenced by Edman degradation of the C fragment are underlined. Panel C also shows an alignment of all published vertebrate *hh* sequences plus some of unpublished sequences from fish and Xenopus. The sequences shown correspond to the region of *Drosophila hh* where the cleavage occurs, and demonstrates the absolute conservation of the Gly-Cys-Phe sequence at the site of cleavage. Panel D shows a SDS-PAGE gel loaded with *in vitro* transcription/translation reactions as described in the previous Examples, using various *hh* genes as templates. dhh is *Drosophila*, *twhh* and zf*shh* are the *twiggy-winkle* and *sonic hh* genes of the zebrafish, and m*shh* is the *shh/Hgh-1/vhh-1*

gene of the mouse. Panel E shows that Edman degradation of the C fragments releases ^{35}S counts on the first but not subsequent rounds for all these proteins, indicating that the site of autoproteolytic cleavage for all of these *hh* proteins is the amide bond to the amino-terminal side of the Cys residue that forms the center of the conserved Gly-Cys-Phe sequence highlighted in panel C.

FIGURE 13 shows the predicted amino acid sequences are shown in single letter code. 13(a) shows sequences common to five distinct *hh*-like genes are shown with a dot indicating identity with the corresponding residue of zebrafish *twiggy-winkle*. 13(b) shows amino acid sequences of *twhh* and *shh* are aligned to those of the *sonic/vhh-1* class from chick and mouse. The amino-terminal hydrophobic stretch common to all four *hh* genes is shaded. The asterisk (*) denotes invariant amino acid residues associated with the proteolytic domain of C fragment from various species. 13(c) shows percent identity of residues carboxy-terminal to the hydrophobic region.

FIGURE 14 shows a comparative expression of *twhh*, *shh*, and *pax-2* during zebrafish embryogenesis.

FIGURE 15 shows the effects of ectopic *hh* on zebrafish development. Wild type zebrafish, *Danio rerio*, Ekkwill Waterlife Resources) were maintained at 28.5°C, some embryos were then cultured overnight at RT. Zebrafish embryos were injected at the 1-8 cell stage with *twhh*, *shh*, or *lacZ*RNA and examined at 28 h of development. (a-c) Dorsal view of the midbrain-hindbrain region; anterior is left. (a) *lacZ*. (b) *twhh*. (c) *shh*. (d-f) Frontal optical section of the forebrain region; anterior is up. (d) *lacZ*. (h) *twhh*. (f) *shh*. (g-1) Lateral view of the eye region; anterior is left. (g) *lacZ*. (h) *twhh*. (i) *twhh*.

FIGURE 16 is a table showing the effects of ectopic expression of *shh*, tw*hh* and tw*hh* mutants on zebrafish embryonic development.

FIGURE 17 shows zebrafish *twiggy-winkle hedgehog* derivatives. 17(a) Cartoons of various *twhh* open reading frames. SS (shaded) is the predicted N-terminal signal sequence for secretion of these proteins and encompasses the first 27 amino acids of each open reading frame. The arrow indicates the predicted internal site of auto-proteolytic cleavage. Amino acid residue numbers are according to Figure 13b. The filled triangle denotes the normal termination codon for the *twhh* open reading frame. Construct U_{HA} contains a mutation that blocks auto-proteolysis (the histidine at residue 273 is changed to an alanine; see Lee, J.J., *et al.*, supra.). Construct $U356_{HA}$ contains a stop codon in place of amino acid residue 357 as well as the H273A mutation in U_{HA}. Construct N encodes just the first 200 amino acids of *twhh*. Construct C has had the codons for residues 31-197 deleted. 17(b) shows *in vitro* translation of the expression constructs shown schematically in part a. Constructs were translated *in vitro* in the presence of ^{35}S methionine and analyzed by autoradiography after SDS-PAGE.

FIGURE 18 shows Northern blot analysis of the effect of hedgehog on expression of various neural markers.

FIGURE 19 shows *hh* synergy with naturally occurring neural markers or agents (*e.g.*, XAG-1, XANF-2, Otx-A, En-2, Krox-20, Xlh box-6, NCAM, and EF-1α).

FIGURE 20 shows Northern blot analysis of the effect of hedgehog N or C on various neural markers.

FIGURE 21 shows ΔN-C interferes with X-*bhh* and N-activity in animal cap explants as shown by RT-PCR analysis.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides two novel polypeptides originally derived from a single precursor protein, both of which have distinct structural and functional characteristics. The proteins are derived from a hedgehog protein and can be naturally produced by auto-proteolytic cleavage of the full-length hedgehog protein. Based on evidence provided herein, which indicates that hedgehog precursor protein and the auto-proteolytic products of hedgehog precursor protein are expressed in the floorplate of the ventral midline of the neural tube and notochord, the invention now provides a method for the induction of proliferation or differentiation of neuronal cells associated with or in close proximity to the floorplate and notochord.

In a first embodiment, the invention provides a substantially pure polypeptide characterized by having an amino acid sequence derived from amino terminal amino acids of a hedgehog protein and having at its carboxy terminus, a glycine-cysteine-phenylalanine (G↓CF) cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide. This fragment is denoted the N-terminal fragment or polypeptide or "N", herein. For example, in the case of the *Drosophila* hedgehog, the N fragment includes amino acids 1-257 of hedgehog protein, wherein amino acids 85-257 have a molecular weight of about 19 kD by non-reducing SDS-PAGE (Amino acid residue numbers 1-257 include non-structural features such as signal sequences.). The G↓CF cleavage site in *Drosophila* hedgehog precursor protein occurs at amino acid residues 257-259. Those of skill in the art will be able to identify the G↓CF cleavage site in other hedgehog genes, as the amino acid location will be similar and the site will be specifically recognized by the autoproteolytic activity of the corresponding C fragment.

The N-terminal polypeptide is also characterized by being cell-associated in cells expressing the polypeptide *in vitro*, and being specifically localized in vertebrate or *Drosophila* cells or embryos, for example. In other words, this N-terminal fragment of

hedgehog, remains close to the site of cellular synthesis. The association of N with the cell is a result of the processing event which involves lipophilic modification of the amino terminal domain. (See Figure 1E) This modification is initiated by the action of the carboxy terminal domain, generating a thioester intermediate; the carboxy-terminal domain thus does not act simply as a protease, although cleavage of a peptide bond does ultimately result from its action. In addition, the N fragment binds to heparin agarose *in vitro*.

The N polypeptide of the invention is characterized by having an amino acid sequence derived from amino terminal amino acids of hedgehog protein, *e.g.*, 1-257 in *Drosophila*, wherein amino acids 1-257 have a molecular weight of about 19 kD by non-reducing SDS-PAGE. The N polypeptide includes smaller fragments which retain the functional characteristics of full length N, *e.g.*, bind to heparin. The hedgehog protein from which N is derived includes, but is not limited to *Drosophila*, *Xenopus*, chicken, zebrafish, mouse, and human. Crystallographic analysis shows the structure of SHH-N includes the presence of a zinc ion. While not wanting to be bound by a particular theory, the presence of the zinc ion is suggestive of zinc hydrolase activity. Zinc hydrolases include proteases such as carboxypeptidase A and thermolysin, lipases such as phospholipase C, and other enzymes such as carbonic anhydrase. Alterations in the zinc hydrolase site of the amino terminal signaling domain may be useful for modulating the range of diffusion of a hedgehog protein or to alter the signaling characteristics of the amino terminal signaling domain. For example, a mutation in the zinc hydrolase site may result in a tethered protein where ordinarily the protein is secreted at a distance. The result would be induction of a cell type not typically induced. Alteration in the zinc site may result in a molecule capable of inducing motor neurons and not floor plate, and vice versa.

The identification of a cell-surface, or extracellular matrix localization of N and its expression in notochord and floor plate-associated cells, provides a means for isolation or specific selection of cells expressing N, *e.g.*, to isolate a notochord sample or to isolate

floor plate cells. In addition, antibodies directed to N are useful for histological analysis of tissues suspected of expressing N protein.

The invention also provides a substantially pure polypeptide characterized by having an amino acid sequence derived from carboxy terminal amino acids of a hedgehog protein and having at its amino terminus a G↓CF cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide. This fragment is denoted the C-terminal fragment or polypeptide or "C", herein. For example, in Drosophila this "C" polypeptide derives from the C-terminal domain of hedgehog precursor protein beginning at amino acid residue 258, wherein the full length C-terminal domain has a molecular weight of about 25 kD by non-reducing SDS-PAGE, a histidine residue at position 72, and has protease activity. The G↓CF cleavage site specifically recognized by the proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide is located at amino acid residues 257-259. As described above for the N fragment, now that the present invention has shown the precise cleavage recognition site for the autoproteolytic domain of hedgehog, those of skill in the art can readily discern the cleavage site in other hedgehog proteins thereby allowing the ready identification of any N or C polypeptide of any hedgehog precursor protein.

The "C" polypeptide of the invention is derived from the C-terminus of a hedgehog precursor protein, beginning at the autoproteolytic cleavage site identified at the GCF amino acid sequence, which in *Drosophila* corresponds to amino acids 257-259. In *Drosophila* the histidine residue found invariably at amino acid residue 329 of the native hedgehog protein, and at amino acid residue 72 of the C polypeptide, is essential for auto-proteolytic cleavage between amino acids 257 and 258 (G and C). Corresponding C-polypeptides of the invention will likewise contain a similarly located histidine residue which can be readily identified, such as by comparison to the *Drosophila* C-polypeptide. Among various species, the proteolytic domain can be characterized by the amino acid sequence -XTXXHLXX-.

The C polypeptide of the invention, unlike N, does not significantly bind to heparin agarose. C is characterized by being released into the culture supernatant of cells expressing C polypeptide *in vitro* and by being localized diffusely in cells and embryos. Because C polypeptide diffuses freely, it would be detectable in various body fluids and tissues in a subject. Identification of C polypeptide expression near the midline of the neural tube, as described herein, provides a useful assay for neural tube closure in an embryo/fetus, for example. The presence of C polypeptide in amniotic fluid would be diagnostic of a disorder in which the neural tube may be malformed.

Altered levels of C polypeptide in cerebrospinal fluid may be indicative of neurodegenerative disorders, for example. Because C polypeptide is released from the cell after synthesis and autoproteolysis of native hedgehog precursor polypeptide, tumors synthesizing and releasing high levels of C polypeptide would be detectable without prior knowledge of the exact location of the tumor.

C fragment is effective in inducing genes of the pituitary and anterior brain as well. In particular, induction is increased by the addition of a member of the TGF-β family of growth factors. For example, human activin in combination with C fragment may be effective in enhancing pituitary cell growth and activity or development.

C fragment is effective in inducing posterior markers of the brain by inhibiting N. Such a fragment is exemplified in Example 18 as ΔN-C. Therefore in another embodiment, the invention includes a polypeptide deleting amino acid residues 28-194 of X-*bhh*. (Autoproteolysis gives a C domain of 198-409 as well as a seven amino acid peptide, representing aa 24-27 and 195-197). This polypeptide blocks the activity of X-*bhh* and N in explants and reduces dorsoanterior structures in embryos. Also included are polynucleotide sequences encoding ΔN-C. ΔN-C is useful for increasing expression of posterior neural markers (*e.g.*, En-2, Krox-20, Xlttbox-6) and decreasing expression of anterior neural markers (*e.g.*, XANF-2, XAG-1, Otx-A) when desirable to do so to modulate neural patterning.

The term "substantially pure" as used herein refers to hedgehog N or C polypeptide which is substantially free of other proteins, lipids, carbohydrates, nucleic acids or other materials with which it is naturally associated. One skilled in the art can purify hedgehog N or C polypeptide using standard techniques for protein purification. The substantially pure polypeptide will yield a single major band on a non-reducing polyacrylamide gel. The purity of the hedgehog N or C polypeptide can also be determined by amino-terminal amino acid sequence analysis.

The invention includes a functional N or C polypeptide, and functional fragments thereof. As used herein, the term "functional polypeptide" or "functional fragment" refers to a polypeptide which possesses a biological function or activity which is identified through a defined functional assay and which is associated with a particular biologic, morphologic, or phenotypic alteration in the cell. Functional fragments of the hedgehog N or C polypeptide include fragments of N or C polypeptide as long as the activity, *e.g.*, proteolytic activity of C polypeptide remains. Smaller peptides containing the biological activity of N or C polypeptide are therefore included in the invention. The biological function, for example, can vary from a polypeptide fragment as small as an epitope to which an antibody molecule can bind to a large polypeptide which is capable of participating in the characteristic induction or programming of phenotypic changes within a cell. A "functional polynucleotide" denotes a polynucleotide which encodes a functional polypeptide as described herein.

Minor modifications of the N or C polypeptide primary amino acid sequence may result in polypeptides which have substantially equivalent activity as compared to the N or C polypeptide described herein. Such modifications may be deliberate, as by site-directed mutagenesis, or may be spontaneous. All of the polypeptides produced by these modifications are included herein as long as the proteolytic activity of C polypeptide, for example, is present. Further, deletion of one or more amino acids can also result in a modification of the structure of the resultant molecule without significantly altering its activity. This can lead to the development of a smaller active molecule which would have

broader utility. For example, it is possible to remove amino or carboxy terminal amino acids which may not be required for N or C polypeptide activity.

The N or C polypeptide of the invention also includes conservative variations of the polypeptide sequence. The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid provided that antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide.

The N fragment of the invention includes both the active form of the polypeptide and the N fragment including the uncleaved signal sequence. For example, in *Drosophila* where the signal sequence is internal (at about amino acids 60-80), the entire uncleaved N fragment beginning at the initiating methionine is included in the invention. Those of skill in the art can readily ascertain the nature and location of the signal sequence by using, for example, the algorithm described in von Heijne, G., *Nucl. Acids Res.* 14:4683, (1986).

The invention also provides an isolated polynucleotide sequence consisting essentially of a polynucleotide sequence encoding a polypeptide having the amino acid sequence of N or C polypeptide of the invention. The term "isolated" as used herein includes polynucleotides substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which it is naturally associated. Polynucleotide sequences of the invention include DNA, cDNA and RNA sequences which encode N or C polypeptide. It is understood that all polynucleotides encoding all or a portion of N or C polypeptide are also included herein, as long as they encode a polypeptide with N or C polypeptide

activity. Such polynucleotides include naturally occurring, synthetic, and intentionally manipulated polynucleotides. For example, N or C polypeptide polynucleotide may be subjected to site-directed mutagenesis. The polynucleotide sequence for N or C polypeptide also includes antisense sequences. The polynucleotides of the invention include sequences that are degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included in the invention as long as the amino acid sequence of N or C polypeptide polypeptide encoded by the nucleotide sequence is functionally unchanged. In addition, the invention also includes a polynucleotide consisting essentially of a polynucleotide sequence encoding a polypeptide having an amino acid sequence of N or C and having at least one epitope for an antibody immunoreactive with N or C polypeptide.

The polynucleotide encoding N or C polypeptide includes the entire polypeptide or fragments thereof, as well as nucleic acid sequences complementary to that sequence. A complementary sequence may include an antisense nucleotide. When the sequence is RNA, the deoxynucleotides A, G, C, and T are replaced by ribonucleotides A, G, C, and U, respectively. Also included in the invention are fragments of the above-described nucleic acid sequences that are at least 15 bases in length, which is sufficient to permit the fragment to selectively hybridize to DNA that encodes the protein under physiological conditions. Preferably, the fragments hybridize under stringent conditions.

DNA sequences of the invention can be obtained by several methods. For example, the DNA can be isolated using hybridization techniques which are well known in the art. These include, but are not limited to: 1) hybridization of genomic or cDNA libraries with probes to detect homologous nucleotide sequences; 2) antibody screening of expression libraries to detect cloned DNA fragments with shared structural features; and 3) PCR amplification of a desired nucleotide sequence using oligonucleotide primers.

Preferably the hedgehog, N, or C polynucleotide of the invention is derived from a vertebrate organism, and most preferably from human. Screening procedures which rely on nucleic acid hybridization make it possible to isolate any gene sequence from any organism, provided the appropriate probe is available. Oligonucleotide probes, which correspond to a part of the sequence encoding the protein in question, can be synthesized chemically. This requires that short, oligopeptide stretches of amino acid sequence must be known. The DNA sequence encoding the protein can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. It is possible to perform a mixed addition reaction when the sequence is degenerate. This includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA. Hybridization is particularly useful in the detection of cDNA clones derived from sources where an extremely low amount of mRNA sequences relating to the polypeptide of interest are present. In other words, by using stringent hybridization conditions directed to avoid non-specific binding, it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the hybridization of the target DNA to that single probe in the mixture which is its complete complement (Wallace, *et al.*, *Nucl. Acid Res.*, 9:879, 1981).

The development of specific DNA sequences encoding hedgehog can also be obtained by: 1) isolation of double-stranded DNA sequences from the genomic DNA; 2) chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and 3) *in vitro* synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a eukaryotic donor cell. In the latter case, a double-stranded DNA complement of mRNA is eventually formed which is generally referred to as cDNA.

Of the three above-noted methods for developing specific DNA sequences for use in recombinant procedures, the isolation of genomic DNA isolates is the least common.

This is especially true when it is desirable to obtain the microbial expression of mammalian polypeptides due to the presence of introns.

The synthesis of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct synthesis of DNA sequences is not possible and the method of choice is the synthesis of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid- or phage-carrying cDNA libraries which are derived from reverse transcription of mRNA which is abundant in donor cells that have a high level of genetic expression. When used in combination with polymerase chain reaction technology, even rare expression products can be cloned. In those cases where significant portions of the amino acid sequence of the polypeptide are known, the production of labeled single or double-stranded DNA or RNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single-stranded form (Jay, *et al.*, *Nucl. Acid Res.*, 11:2325, 1983).

A preferred method for obtaining genomic DNA, for example, is Polymerase Chain Reaction (PCR), which relies on an *in vitro* method of nucleic acid synthesis by which a particular segment of DNA is specifically replicated. Two oligonucleotide primers that flank the DNA fragment to be amplified are utilized in repeated cycles of heat denaturation of the DNA, annealing of the primers to their complementary sequences, and extension of the annealed primers with DNA polymerase. These primers hybridize to opposite strands of the target sequence and are oriented so that DNA synthesis by the polymerase proceeds across the region between the primers. Since the extension products themselves are also complementary to and capable of binding primers, successive cycles of amplification essentially double the amount of the target DNA synthesized in the previous cycle. The result is an exponential accumulation of the specific target fragment, approximately 2^n, where n is the number of cycles of amplification performed (see PCR

Protocols, Eds. Innis, *et al.*, Academic Press, Inc., 1990, incorporated herein by reference).

A cDNA expression library, such as λgt11, can be screened indirectly for hedgehog, N, or C polypeptides having at least one epitope, using antibodies specific for hedgehog, N, or C. Such antibodies can be either polyclonally or monoclonally derived and used to detect expression product indicative of the presence of the desired hedgehog cDNA.

The polynucleotide sequence for hedgehog, N, or C, also includes sequences complementary to the polynucleotide encoding hedgehog, N or C (antisense sequences). Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (Weintraub, *Scientific American*, 262:40, 1990). The invention embraces all antisense polynucleotides capable of inhibiting production of hedgehog, N, or C polypeptide. In the cell, the antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule. The antisense nucleic acids interfere with the translation of the mRNA since the cell will not translate a mRNA that is double-stranded. Antisense oligomers of about 15 nucleotides are preferred, since they are easily synthesized and are less likely to cause problems than larger molecules when introduced into the target hedgehog, N, or C-producing cell. The use of antisense methods to inhibit the translation of genes is well known in the art (Marcus-Sakura, *Anal. Biochem.*, 172:289, 1988). Inhibition of target nucleotide would be desirable, for example, in inhibiting cell-proliferative disorders, such as certain tumors, which are mediated by hedgehog, N or C.

In addition, ribozyme nucleotide sequences for hedgehog, N or C are included in the invention. Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences which encode these RNAs, it is possible to engineer molecules that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, *J. Amer. Med. Assn.*, 260:3030, 1988). A major advantage

of this approach is that, because they are sequence-specific, only mRNAs with particular sequences are inactivated.

There are two basic types of ribozymes namely, *tetrahymena*-type (Hasselhoff, *Nature*, 334:585, 1988) and "hammerhead"-type. *Tetrahymena*-type ribozymes recognize sequences which are four bases in length, while "hammerhead"-type ribozymes recognize base sequences 11-18 bases in length. The longer the recognition sequence, the greater the likelihood that sequence will occur exclusively in the target mRNA species. Consequently, hammerhead-type ribozymes are preferable to *tetrahymena*-type ribozymes for inactivating a specific mRNA species and 18-based recognition sequences are preferable to shorter recognition sequences.

DNA sequences encoding hedgehog, N or C can be expressed *in vitro* by DNA transfer into a suitable host cell. "Host cells" are cells in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However, such progeny are included when the term "host cell" is used. Methods of stable transfer, meaning that the foreign DNA is continuously maintained in the host, are known in the art.

In the present invention, the hedgehog, N or C polynucleotide sequences may be inserted into a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of the hedgehog, N or C genetic sequences. Such expression vectors contain a promoter sequence which facilitates the efficient transcription of the inserted genetic sequence of the host. The expression vector typically contains an origin of replication, a promoter, as well as specific genes which allow phenotypic selection of the transformed cells. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in bacteria (Rosenberg, *et al.*, *Gene*, 56:125, 1987), the pMSXND expression vector for expression in mammalian

cells (Lee and Nathans, *J. Biol. Chem.*, 263:3521, 1988) and baculovirus-derived vectors for expression in insect cells. The DNA segment can be present in the vector operably linked to regulatory elements, for example, a promoter (e.g., T7, metallothionein I, or polyhedrin promoters).

Polynucleotide sequences encoding hedgehog, N or C can be expressed in either prokaryotes or eukaryotes, although post-translational modification of eukaryotically derived polypeptides, such as carboxylation, would occur in a eukaryotic host. Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Biologically functional viral and plasmid DNA vectors capable of expression and replication in a host are known in the art. Such vectors are used to incorporate DNA sequences of the invention.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing the hedgehog, N or C coding sequence and appropriate transcriptional/translational control signals. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* recombination/genetic techniques. See, for example, the techniques described in Maniatis, *et al.*, 1989 Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y.

A variety of host-expression vector systems may be utilized to express the hedgehog, N or C coding sequence. These include but are not limited to microorganisms such as bacteria transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing the hedgehog, N or C coding sequence; yeast transformed with recombinant yeast expression vectors containing the hedgehog, N or C coding sequence; plant cell systems infected with recombinant virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (*e.g.*, Ti plasmid) containing the Hedgehog, N or C coding sequence; insect cell systems infected with recombinant virus expression

vectors (*e.g.*, baculovirus) containing the hedgehog, N or C coding sequence; or animal cell systems infected with recombinant virus expression vectors (*e.g.*, retroviruses, adenovirus, vaccinia virus) containing the hedgehog, N or C coding sequence, or transformed animal cell systems engineered for stable expression.

Depending on the host/vector system utilized, any of a number of suitable transcription and translation elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, etc., may be used in the expression vector (see *e.g.*, Bitter, *et al.*, 1987, Methods in Enzymology, 153:516-544). For example, when cloning in bacterial systems, inducible promoters such as pL of bacteriophage γ, plac, ptrp, ptac (ptrp-lac hybrid promoter) and the like may be used. When cloning in mammalian cell systems, promoters derived from the genome of mammalian cells (*e.g.*, metallothionein promoter) or from mammalian viruses (*e.g.*, the retrovirus long terminal repeat; the adenovirus late promoter; the vaccinia virus 7.5K promoter) may be used. Promoters produced by recombinant DNA or synthetic techniques may also be used to provide for transcription of the inserted hedgehog, N or C coding sequence.

In bacterial systems a number of expression vectors may be advantageously selected depending upon the use intended for the expressed. For example, when large quantities of hedgehog, N or C are to be produced, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Those which are engineered to contain a cleavage site to aid in recovering are preferred. Such vectors include but are not limited to the *E. coli* expression vector pUR278 (Ruther, *et al.*, *EMBO J.*, 2:1791, 1983), in which the Hedgehog, N or C coding sequence may be ligated into the vector in frame with the lac Z coding region so that a hybrid -lac Z protein is produced; pIN vectors (Inouye and Inouye, *Nucleic Acids Res.*, 13:3101, 1985; Van Heeke and Schuster, *J. Biol. Chem.* 264:5503, 1989) and the like.

In yeast, a number of vectors containing constitutive r inducible promoters may be used. For a review see, Current Protocols in Molecular Biology, Vol. 2, 1988, Ed. Ausubel, *et*

al., Greene Publish. Assoc. & Wiley Interscience, Ch. 13; Grant, *et al.*, 1987, Expression and Secretion Vectors for Yeast, in Methods in Enzymology, Eds. Wu and Grossman, 31987, Acad. Press, N.Y., Vol. 153, pp.516-544; Glover, 1986, DNA Cloning, Vol. II, IRL Press, Wash., D.C., Ch. 3; and Bitter, 1987, Heterologous Gene Expression in Yeast, Methods in Enzymology, Eds. Berger and Kimmel, Acad. Press, N.Y., Vol. 152, pp. 673-684; and The Molecular Biology of the Yeast Saccharomyces, 1982, Eds. Strathern, *et al.*, Cold Spring Harbor Press, Vols. I and II. A constitutive yeast promoter such as ADH or LEU2 or an inducible promoter such as GAL may be used (Cloning in Yeast, Ch. 3, R. Rothstein In: DNA Cloning Vol.11, A Practical Approach, Ed. DM Glover, 1986, IRL Press, Wash., D.C.). Alternatively, vectors may be used which promote integration of foreign DNA sequences into the yeast chromosome.

In cases where plant expression vectors are used, the expression of the hedgehog, N or C coding sequence may be driven by any of a number of promoters. For example, viral promoters such as the 35S RNA and 19S RNA promoters of CaMV (Brisson, *et al.*, *Nature*, 310:511, 1984), or the coat protein promoter to TMV (Takamatsu, *et al.*, *EMBO J.*, 6:307, 1987) may be used; alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi, *et al.*, *EMBO J.*, 3:1671-1680, 1984; Broglie, *et al.*, *Science*, 224:838, 1984); or heat shock promoters, *e.g.*, soybean hsp17.5-E or hsp17.3-B (Gurley, *et al.*, *Mol. Cell. Biol.*, 6:559, 1986) may be used. These constructs can be introduced into plant cells using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, microinjection, electroporation, etc. For reviews of such techniques see, for example, Weissbach and Weissbach, 1988, Methods for Plant Molecular Biology, Academic Press, NY, Section VIII, pp. 421-463; and Grierson and Corey, 1988, Plant Molecular Biology, 2d Ed., Blackie, London, Ch. 7-9.

An alternative expression system which could be used to express is an insect system. In one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The hedgehog, N or C coding sequence may be cloned into non-essential regions (for example

the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the hedgehog, N or C coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (*i.e.*, virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed. (*e.g.*, see Smith, *et al.*, *J. Viol.*, 46:584, 1983; Smith, U.S. Patent No. 4,215,051).

Eukaryotic systems, and preferably mammalian expression systems, allow for proper post-translational modifications of expressed mammalian proteins to occur. Eukaryotic cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, phosphorylation, and advantageously, secretion of the gene product may be used as host cells for the expression of hedgehog, N or C. Mammalian cell lines may be preferable. Such host cell lines may include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, -293, and WI38.

Mammalian cell systems which utilize recombinant viruses or viral elements to direct expression may be engineered. For example, when using adenovirus expression vectors, the hedgehog, N or C coding sequence may be ligated to an adenovirus transcription/-translation control complex, *e.g.*, the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (*e.g.*, region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the protein in infected hosts (*e.g.*, see Logan and Shenk, *Proc. Natl. Acad. Sci. USA*, 81:3655, 1984). Alternatively, the vaccinia virus 7.5K promoter may be used. (*e.g.*, see, Mackett, *et al.*, *Proc. Natl. Acad. Sci. USA*, 79:7415, 1982; Mackett, *et al.*, *J. Virol.*, 49: 857, 1984; Panicali, *et al.*, *Proc. Natl. Acad. Sci. USA*, 79:4927, 1982). Of particular interest are vectors based on bovine papilloma virus which have the ability to replicate as extrachromosomal elements (Sarver, *et al.*, *Mol. Cell. Biol.*, 1:486, 1981). Shortly after entry of this DNA into mouse cells, the plasmid replicates to about 100 to 200 copies per

cell. Transcription of the inserted cDNA does not require integration of the plasmid into the host's chromosome, thereby yielding a high level of expression. These vectors can be used for stable expression by including a selectable marker in the plasmid, such as, for example, the neo gene. Alternatively, the retroviral genome can be modified for use as a vector capable of introducing and directing the expression of the hedgehog, N or C gene in host cells (Cone and Mulligan, *Proc. Natl. Acad. Sci. USA*, 81:6349, 1984). High level expression may also be achieved using inducible promoters, including, but not limited to, the metallothionine IIA promoter and heat shock promoters.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with the hedgehog, N or C cDNA controlled by appropriate expression control elements (*e.g.*, promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. For example, following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler, *et al.*, *Cell*, 11: 223, 1977), hypoxanthine-guanine phosphoribosyltransferase (Szybalska and Szybalski, *Proc. Natl. Acad. Sci. USA*, 48:2026, 1962), and adenine phosphoribosyl-transferase (Lowy, *et al.*, *Cell*, 22: 817, 1980) genes can be employed in tk^-, $hgprt^-$ or $aprt^-$ cells respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, *et al.*, *Natl. Acad. Sci. USA*, 77: 3567, 1980; O'Hare, *et al.*, *Proc. Natl. Acad. Sci. USA*, 78: 1527, 1981); gpt, which confers resistance to mycophenolic acid (Mulligan and Berg, *Proc. Natl. Acad. Sci. USA*, 78: 2072, 1981; neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, *et al.*, *J. Mol. Biol.*, 150:1, 1981); and hygro, which confers resistance to hygromycin (Santerre, *et al.*, *Gene*, 30:147, 1984) genes. Recently, additional

selectable genes have been described, namely trpB, which allows cells to utilize indole in place of tryptophan; hisD, which allows cells to utilize histinol in place of histidine (Hartman and Mulligan, *Proc. Natl. Acad. Sci. USA*, 85:8047, 1988); and ODC (ornithine decarboxylase) which confers resistance to the ornithine decarboxylase inhibitor, 2-(difluoromethyl)-DL-ornithine, DFMO (McConlogue L., 1987, In: Current Communications in Molecular Biology, Cold Spring Harbor Laboratory ed.).

Transformation of a host cell with recombinant DNA may be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *E. coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method using procedures well known in the art. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell if desired.

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used. Eukaryotic cells can also be cotransformed with DNA sequences encoding the hedgehog, N or C of the invention, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein. (see for example, *Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982).

Isolation and purification of microbial expressed polypeptide, or fragments thereof, provided by the invention, may be carried out by conventional means including preparative chromatography and immunological separations involving monoclonal or polyclonal antibodies.

The invention includes antibodies immunoreactive with or which bind to hedgehog, N or C polypeptide or functional fragments thereof. Antibody which consists essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations are provided. Monoclonal antibodies are made from antigen containing fragments of the protein by methods well known to those skilled in the art (Kohler, *et al.*, *Nature*, 256:495, 1975). The term antibody as used in this invention is meant to include intact molecules as well as fragments thereof, such as Fab and $F(ab')_2$, which are capable of binding an epitopic determinant on hedgehog, N or C. The antibodies of the invention include antibodies which bind to the N or C polypeptide and which bind with immunoreactive fragments N or C.

The term "antibody" as used in this invention includes intact molecules as well as fragments thereof, such as Fab, $F(ab')_2$, and Fv which are capable of binding the epitopic determinant. These antibody fragments retain some ability to selectively bind with its antigen or receptor and are defined as follows:

(1) Fab, the fragment which contains a monovalent antigen-binding fragment of an antibody molecule can be produced by digestion of whole antibody with the enzyme papain to yield an intact light chain and a portion of one heavy chain;

(2) Fab', the fragment of an antibody molecule can be obtained by treating whole antibody with pepsin, followed by reduction, to yield an intact light chain and a portion of the heavy chain; two Fab' fragments are obtained per antibody molecule;

(3) $(Fab')_2$, the fragment of the antibody that can be obtained by treating whole antibody with the enzyme pepsin without subsequent reduction; $F(ab')_2$ is a dimer of two Fab' fragments held together by two disulfide bonds;

(4) Fv, defined as a genetically engineered fragment containing the variable genetically fused single chain molecule.

Methods of making these fragments are known in the art. (See for example, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1988), incorporated herein by reference).

As used in this invention, the term "epitope" means any antigenic determinant on an antigen to which the paratope of an antibody binds. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three dimensional structural characteristics, as well as specific charge characteristics.

Antibodies which bind to the hedgehog, N or C polypeptide of the invention can be prepared using an intact polypeptide or fragments containing small peptides of interest as the immunizing antigen. The polypeptide such as N or C, or fragments thereof used to immunize an animal can be derived from translated cDNA or chemical synthesis which can be conjugated to a carrier protein, if desired. Such commonly used carriers which are chemically coupled to the peptide include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

If desired, polyclonal or monoclonal antibodies can be further purified, for example, by binding to and elution from a matrix to which the polypeptide or a peptide to which the antibodies were raised is bound. Those of skill in the art will know of various techniques common in the immunology arts for purification and/or concentration of polyclonal antibodies, as well as monoclonal antibodies (See for example, Coligan, *et al.*, Unit 9, *Current Protocols in Immunology*, Wiley Interscience, 1991, incorporated by reference). It is also possible to use the anti-idiotype technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to

a first monoclonal antibody will have a binding domain in the hypervariable region which is the "image" of the epitope bound by the first monoclonal antibody.

Antibodies as described herein as having specificity for N polypeptide, *e.g.*, Ab1 (residues 83-160), are useful for specific identification of cells or tissues expressing the N fragment of hedgehog. Similarly, antibodies described herein as having specificity for C polypeptide, *e.g.*, Ab2 (residues 300-391), are useful for specific identification of cells or tissues expressing the C fragment of hedgehog. Both antibodies, naturally, will also detect native hedgehog polypeptide.

The N and C-specific antibodies of the invention are useful for purification of N and C polypeptide, respectively, especially using the antibodies immobilized on solid phase. By contacting a sample with anti-N antibody, both N and native hedgehog polypeptides can be isolated. By next contacting the sample removed by anti-N antibodies, with anti-C antibodies, the native hedgehog polypeptide is removed, thus allowing purification of N polypeptide. In a similar manner, C polypeptide can be antibody purified from a sample.

Monoclonal antibodies of the invention are suited for use, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the monoclonal antibodies in these immunoassays can be detectably labeled in various ways. Examples of types of immunoassays which can utilize monoclonal antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the monoclonal antibodies of the invention can be done utilizing immunoassays which are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

The term "immunometric assay" or "sandwich immunoassay", includes simultaneous sandwich, forward sandwich and reverse sandwich immunoassays. These terms are well understood by those skilled in the art. Those of skill will also appreciate that antibodies according to the present invention will be useful in other variations and forms of assays which are presently known or which may be developed in the future. These are intended to be included within the scope of the present invention.

Monoclonal antibodies can be bound to many different carriers and used to detect the presence of N or C polypeptide. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such using routine experimentation.

For purposes of the invention, N or C polypeptide may be detected by the monoclonal antibodies when present in biological fluids and tissues. Any sample containing a detectable amount of N or C can be used. A sample can be a liquid such as urine, saliva, cerebrospinal fluid, blood, serum and the like, or a solid or semi-solid such as tissues, feces, and the like, or, alternatively, a solid tissue such as those commonly used in histological diagnosis. C polypeptide in particular is detectable in biological samples, since it tends to diffuse more readily than N polypeptide.

In performing the assays it may be desirable to include certain "blockers" in the incubation medium (usually added with the labeled soluble antibody). The "blockers" are added to assure that non-specific proteins, proteases, or anti-heterophilic immunoglobulins to anti-C or N immunoglobulins present in the experimental sample do not cross-link or destroy the antibodies on the solid phase support, or the radiolabeled indicator antibody, to yield false positive or false negative results. The selection of

"blockers" therefore may add substantially to the specificity of the assays described in the present invention.

The invention also provides a method for modulating proliferation or differentiation of neuronal cells comprising contacting the cells with a hedgehog polypeptide. The hedgehog polypeptide may be a native hedgehog polypeptide, or a N or C polypeptide, or functional fragments thereof. Preferably, the modulation is induction of proliferation or differentiation of a particular cell type. This can involve either synergistic positive induction of neuronal cells by N, or negative modulation by delta N-C for example (Lai, *et al.*, *Development* 121:2349, 1995). Delta N-C enhances expression of posterier relative to anterior neural genes and does so through inhibition of N (see EXAMPLE 18 and Figure 18D). In addition to hedgehog polypeptide, a TGF-β factor may also be utilized in the method of the invention.

Previous studies with the rat hedgehog gene showed that co-culture of cells expressing rat hedgehog precursor gene, with explant from neural tube, was sufficient to induce formation of motor neurons and floor plate from the explant (Jessesl, T., and Dodd, J., In *Cell-Cell Signaling in Vertebrate Development* (ed. E.J. Robertson, *et al.*, pp 139-155, San Diego, Ca.), 1993). Therefore, based on the Examples herein showing that hedgehog is expressed near the floorplate of the ventral midline of the neural tube and notochord, neuronal cells substantially derived from floor plate neuronal cells can be induced by contacting the cells with hedgehog, N or C polypeptide. As used herein, the term "substantially derived", refers to those cells from the floor plate or proximate to the floor plate. For example, such cells include motor neurons and dopaminergic neurons. Those of skill in the art will be able to identify other neuronal cells substantially derived from the floor plate. Preferably the cells are vertebrate cells and most preferably, human cells.

In addition, as described herein in the Examples, hedgehog, and particularly C fragment, induces the expression of pituitary genes. Hedgehog is also effective in inducing anterior brain gene expression as exemplified by the OTX-A marker. Further, the addition of a

TGF-β family member, for example activin, may be used to further induce expression of such genes. Other TGF-β family members will be known to those of skill in the art. This apparent synergy of *hh* fragments with TGF-β family members occurs through the TGF-β protein inducing expression of neural inducers such as noggin and follistatin. The *hh* fragment then synergizes with these inducers to pattern neural gene expression.

hh fragments may also be useful as nerve-sparing agents or in restoring or promoting appropriate patterning during the healing of major limb trauma. In addition, the N and C fragments may be useful in the area of genetic counseling. Specifically, familial midline defects such as cyclopia, polydactyly or neural tube defects may be diagnosed by mapping close to *hh*. Since autoproteolytic defects may be responsible for the disorders, N or C therapy could be provided.

The invention also provides an autoproteolytic fusion protein comprising a first polypeptide including the proteolytic domain of the C polypeptide of the invention, a cleavage site recognized by the first polypeptide, and a second polypeptide. (It is understood that the first and second polypeptides can be reversed.) The auto-proteolytic activity of the native hedgehog protein is found entirely within the C polypeptide, therefore, the C polypeptide is useful for producing a fusion polypeptide which can then be cleaved at the junction of the C polypeptide and the second polypeptide. The fusion protein may optionally have a purification tag, such as a poly-histidine tag for isolation on a nickel column, or an antibody epitope tag, preferably on the C fragment. The cleavage site includes the sequence "GCF", which is recognized by the proteolytic domain of the C polypeptide and is utilized to cleave the second polypeptide from the C fragment. Also included in the invention is a polynucleotide encoding the fusion protein of the invention.

The invention also provides a method for producing an autoproteolytic fusion protein comprising operably linking a first polynucleotide, wherein the first polynucleotide encodes a first polypeptide including the proteolytic domain f the C polypeptide of the

invention and the cleavage site recognized by the proteolytic domain, and a second polynucleotide encoding a second polypeptide. As described above, the fusion protein may also include a carrier peptide and/or a purification tag.

The C polypeptide or functional fragment thereof is useful as a fusion partner to cause lipophilic modification and tethering of other proteins *in vivo* or *in vitro*. Such fusion proteins may be desirable for factors whose activity is required in a localized manner, either by targeting DNA constructs to specific cells or by introducing cells transfected with specific DNA constructs, for example. It may be desirable to lipid-modify a normally secreted protein in order to produce a cell-associated protein. For example, it may be desirable to produce a viral antigen that remains cell associated.

Alternatively, the C polypeptide or functional fragments thereof can be used as a fusion partner with a protein of interest (*e.g.*, Protein X fused to hh-C domain). Such fusions form thioesters at the junction between Protein X and hh-C (via an S to N shift). The thioesters are then available as substrates for a peptide ligation reaction in which any peptide or protein having an amino terminal cysteine (Peptide Y) is added and undergoes spontaneous rearrangement (S to N shift) that generates a stable peptide bond between Protein X and Peptide Y (Protein X-peptide bond-Peptide Y). For example, a protein that is toxic when produced *in vivo* could be produced *in vitro* using the hh-C domain fusion protein method.

The fusion polypeptide may also include an optional carrier peptide. The "carrier peptide", or signal sequence, is located at the amino terminal end of the fusion peptide sequence. In the case of eukaryotes, the carrier peptide is believed to function to transport the fusion polypeptide across the endoplasmic reticulum. The secretory protein is then transported through the Golgi apparatus, into secretory vesicles and into the extracellular space or, preferably, the external environment. Carrier peptides which can be utilized according to the invention include pre-pro peptides which contain a proteolytic enzyme recognition site. Acceptable carrier peptides include the amino terminal pro-

region of calcitonin or other hormones, which undergo cleavage at the flanking dibasic sites. However, it should be noted that the invention is not limited to the use of any particular peptide as a carrier. Other carrier peptides are known to those skilled in the art or can be readily ascertained without undue experimentation.

In one embodiment of the invention, a carrier peptide which is a signal sequence is included in the expression vector, specifically located adjacent to the N-terminal end of the fusion polypeptide. This signal sequence allows the fusion protein to be directed toward the endoplasmic reticulum. Typically, the signal sequence consists of a leader of from about 16 to about 29 amino acids, starting with two or three polar residues and continuing with a high content of hydrophobic amino acids; there is otherwise no detectable conservation of sequence known. Such signal sequences are known to those of skill in the art, and include the naturally occurring signal sequence derived from a hedgehog protein.

The fusion polypeptide of the invention includes a polypeptide encoded by a structural gene, preferably at the amino-terminus of the fusion polypeptide. Any structural gene is expressed in conjunction with the C-polypeptide (polynucleotide) and optionally a carrier peptide. The structural gene is operably linked with the carrier in an expression vector so that the fusion polypeptide is expressed as a single unit.

The identification of the autoproteolysis of hedgehog into the N and C domains is useful in a screening method to identify compounds or compositions which affect this processing activity. Thus, in another embodiment, the invention provides a method for identifying a composition which affects hh processing, which can be determined by activity or gene expression, comprising incubating the components, which include the composition to be tested (*e.g.*, a drug, a small molecule, a protein) and a hh polypeptide or a recombinant cell expressing hedgehog or a gene encoding a C domain or functional fragment thereof operably linked to an N domain or functional fragment thereof, under conditions sufficient to allow the components to interact, then subsequently measuring

the effect the composition has on hedgehog activity or expression. Fragments of hedgehog polypeptide or polynucleotide can be used in the method of the invention as long as autoproteolytic activity remains (*e.g.*, the construct exemplified in Figure 12a and 12b, Example 10). The observed effect on hh may be either inhibitory or stimulatory. For example, one can determine whether the N domain is associated with the cell, or whether the N domain is secreted into the medium, in other words, whether incomplete processing has occurred. Such methods for determining the effect of the compound or composition on hh processing include those described herein (see Example 10, Figure 12a and 12b) such as time course of autoproteolytic cleavage or course of cleavage based on concentration ranges. Alternatively, the effect of the composition on hh can be determined by the expression of anterior or posterior neural markers. Other methods for determining the effect of a composition on processing of N and C will be known to those of skill in the art. Various labels can be used to detect the N and C domains, for example, a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme could be used. Those of ordinary skill in the art will know of other suitable labels or will be able to ascertain such, using routine experimentation.

As used herein, "hh activity" as described in the screening method refers preferably to autoproteolytic activity. However, it is understood, that one of skill in the art could use the above-described screening assay to identify a composition having an affect on other hh activities, for example, zinc hydrolase activity. Appropriate assays for determining the effect on such activities will be known to those of skill in the art.

The following examples are intended to illustrate but not limit the invention. While they are typical of those that might be used, other procedures known to those skilled in the art may alternatively be used.

EXAMPLE 1
HEDGEHOG PROTEIN PROCESSING

The full length form of the *hh* protein (F) migrates with a mobility corresponding to a relative molecular mass of 46 kD. FIGURES 1 (A) and (C) are immunoblots with antibodies against amino- (Ab1) and carboxy-terminal (Ab2) epitopes. GST fusion proteins containing either residues 83 to 160 or 300 to 391 from HH protein were expressed in *Escherichia coli*, purified as recommended [F. M. Ausubel, *et al.*, *Current Protocols in Molecular Biology* (Greene and Wiley-Interscience, New York, 1991)], and used to immunize rabbits by standard methods. The antibodies were affinity purified on a column of His_6-U protein [E. Harlow and D. Lane, *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988)] linked to Affi-Gel 10 beads (Bio-Rad). The purification was performed as described (Harlow and Lane, *supra*) except that the acid and base elutions contained 10 percent dioxane. Biotinylated *hh* antibodies were prepared by purifying the rabbit antisera over a protein A column, followed by biotinylation with the use of the Immunoprobe biotinylation kit (Sigma). Immunoprecipitations were performed as described [Harlow and Lane] with the use of cold RIPA lysis buffer containing 0.25 mM phenylmethylsulfonyl fluoride (PMSF) and 5 mM EDTA for tissue homogenization. Lysates were precleared twice with pre-immune rabbit serum plus protein A beads (Gibco-BRL). Affinity-purified antibodies or preimmune serum was then added, and the immunoprecipitation was performed with protein A beads, with the use of NP-40 lysis buffer for the washes.

Immunoblots were performed with affinity purified Ab1 or Ab2 by either of two chemiluminescence based protocols. In the first protocol (used in Figures 1, 3, and 5) samples were resolved on 15 percent or 12 percent SDS-polyacrylamide gels (F. M. Ausubel et al., *supra*) and transferred to Magnagraph nylon membranes (MSI) by electroblotting. Blots were developed with the use of an alkaline phosphatase conjugated donkey anti-rabbit IgG secondary antibody and Lumi-Phos 530 (Boehringer Mannheim) under recommended conditions. In the second protocol (used in FIGURE 8), samples

were transferred to nitrocellulose filters (Schleicher and Schuell), and blots were developed using ECL reagents (Amersham) as recommended. The secondary antibody in this case was horseradish peroxidase conjugated goat anti-rabbit IgG (Jackson ImmunoResearch). Lanes contain protein from induced untransfected S2 cells (lanes 1 and 13), transfected S2 cells induced to express *hh* (lanes 2 and 14), imaginal discs (lanes 3 and 15), wild type embryos (lanes 6 and 18), and *in vitro* translations of synthetic h mRNA both in the presence (lanes 5 and 17) and absence of microsomes (lanes 4 and 16).

cDNAs encoding various *hh* protein species were cloned into the pMK33 vector, which allows for inducible expression under metallothionein promoter control (M. R. Koelle et al., *Cell* 67:59,1991). Stable S2 cell lines were made by transfection of the *hh*/pMK33 plasmids with constant selection for hygromycin resistance. Proteins were expressed by plating a log phase culture of cells diluted to 0.1 A_{595} units, waiting 48 hours, inducing with $CuSO_4$ at 0.2 mM final concentration, and harvesting the cells and/or supernatant 24 hours later. Cell samples for immunoblotting were made by adding 10 volumes of 1X SDS PAGE loading buffer to pelleted cells.

In vitro translations were performed with the use of the TNT coupled transcription-translation system (Promega). ^{35}S methionine (DuPont NEN) was used for detection by autoradiography. In the heparin binding experiment *in vitro* translation lysate with microsomes that produce wild-type *hh* protein was added to heparin agarose (Sigma) or Sepharose CL-4B (Pharmacia) beads pre-equilibrated with heparin binding buffer (HBB; 20 mM Tris (7.4), 150 mM NaCl, 0.1 percent Triton X-100). Samples were incubated at 4° C for four hours with gentle rocking. After pelleting the beads, supernatants in some samples were analyzed (lanes 2 and 4). The beads were then washed 5 times with chilled HBB and samples (lanes 3 and 5) were subsequently eluted at 80° C for 10 minutes in SDS PAGE loading buffer (F. M. Ausubel et al., *supra*).

Embryos from the wild-type Canton-S line and from the matings, hs*hh*/hs*hh* or hs*hh* H329A/hs*hh* H329A X *y*, *Sco/CyO, enlacZ11::wg* (Kassis, *et al., Proc. Natl. Acad. Sci.*

U.S.A. 89: 1919, 1992), were collected 0 to 16 hours after egg laying (AEL) at 25° C. They were heat shocked for 30 minutes at 37° C and allowed to recover for 1 hour at 25° C. Embryos in FIGURE 1 (Canton-S) were collected 4 to 8 hours AEL at 25° C. In preparation for immunoblotting, all embryos were dechorionated in 2.6 percent sodium hypochlorite and homogenized in 10 volumes of 1X SDS PAGE loading buffer.

Multiple species were detected and minor cross reactive bands are seen in most samples including extracts of induced untransfected S2 cells (lanes 1 and 13). One of these bands (occurring in both panels) co-migrates with U (at 39 kD) and is particularly abundant in lane 6 of FIGURE 1 (A).

FIGURES 1 (B) and (D) are blots of samples immunoprecipitated with Ab1 (B, lanes 7-9), Ab2 (D, lanes 19-21), or pre-immune serum (B, lanes 10-12 and D, lanes 22-24). Detection was with biotinylated derivatives of Ab1 (B) and Ab2 (D). Samples used were: induced untransfected S2 cells, lanes 7, 10, 19 and 22; transfected S2 cells induced to express *hh*, lanes 8, 11, 20 and 23; and embryos, lanes 9, 12, 21 and 24. For either antibody, *hh* protein fragments were specifically immunoprecipitated from *hh* expressing cells and embryos, but not from untransfected cells. (E) In the schematic diagram, cleavage sites are denoted by arrows. The cleavage site marked by the asterisk is inferred by identification of only one cleavage product and may therefore occur at another location within the C fragment. The first two columns to the right of the diagram indicate the reactivity of Ab1 and Ab2 to each *hh* fragment. The other columns indicate the presence (+) or absence (-) of each *hh* fragment in the various samples. Parentheses around F and N_{ss} indicate that these species are detected in *in vitro* translation reactions but not *in vivo*.

The 46kD species was detected from *in vitro* translation extracts by Ab1 and Ab2 (FIGURE 1, lanes 4 and 16), and was partially converted to a species of 39 kD (U) when translation occurred in the presence of microsomes (FIGURE 1, lanes 5 and 17). A 39 kD species co-migrating with U is also present in extracts from all *in vivo* sources, but

none of these extracts contain detectable levels of F. U represents the signal-cleaved form of F; signal cleavage thus appears to be relatively inefficient *in vitro*, as reported previously, (J. J. Lee, *et al.*, *Cell*, 71:33, 1992), but is highly efficient *in vivo*. To confirm that signal cleavage indeed is occurring at this unusual internal location, a mutation that changes residue S_{94} to N at the predicted signal cleavage site was introduced. This mutation prevented conversion by microsomes of F to U and also produced a species that comigrated with F upon transfection into cultured S2 cells. The effects of independently mutating the two methionine codons present upstream of the signal sequence were also examined. *In vitro* translation of the sequence in which the first methionine is removed produces a protein species intermediate in mobility between F and U, and this species is converted to a species that comigrates with U in the presence of microsomes or when produced *in vivo*. Alteration of the second methionine codon caused no change in the electrophoretic mobility of *Hh* protein produced *in vivo* or *in vitro*.

Smaller species of *Hh* proteins from *in vivo* sources have been reported previously (T. Tabata and T. B. Kornberg, *Cell* 76: 89, 1994). The latter study examined not endogenous proteins, but proteins induced to express at high levels from exogenously introduced constructs. The antibody used did not distinguish epitopes from distinct portions of the molecule.

In addition to signal cleavage, a further cleavage of the U precursor is responsible for generating other forms of *hh* protein observed *in vivo*. This was deduced from the observation that Ab1 and Ab2 both detected the U (uncleaved) species, but also interacted individually with smaller protein species expressed endogenously in embryos and imaginal discs or with species expressed upon introduction of the *hh* gene into S2 cells. Ab1 thus interacts with a 19kD species from all of these tissues (FIGURE 1, lanes 2, 3, 6, 8, 9), while Ab2 interacts with a 25 kD species and a 16 kD species (FIGURE 1, lanes 14, 15, 18, 20, 21). The 19 kD species hereafter is referred to as N (N-terminal fragment), the 25 kD species as C (C-terminal fragment) and the 16 kD species as C*; these species represent the major forms of endogenous *hh* protein present *in vivo*.

The proposed cleavages by which these species arise are shown schematically in the bottom portion of FIGURE 1. The N and C species are uniquely detected by Ab1 and Ab2, respectively, and the sum of the relative masses of the two smaller species is roughly equivalent to the relative mass of U. The electrophoretic mobilities of the F and U species are somewhat at variance with their predicted relative masses (52.1 kD and 43.3 kD, respectively). The identities of these species were confirmed by *in vitro* translation of a variety of *hedgehog* open reading frames modified to contain different extents of sequence at the NH_2- or COOH- terminus, and by insertion of epitope tags. The migration anomalies appear to be associated with protein species in which sequences from both the NH_2- and COOH-terminal fragments are simultaneously present. The mobilities of the NH_2- and COOH-terminal fragments, in contrast, correspond to relative masses (19 kD and 25 kD, respectively) that sum to yield 44 kD, roughly equivalent to the expected relative mass of U.

A simple mechanism that could account for the derivation of the two smaller species therefore would be a single internal cleavage of the U precursor. Processing of the *hh* protein when translated *in vitro* also yields a 25 kD species (C; lanes 16 and 17) and either a 29 kD or 19 kD (N) species (lanes 4 and 5). The 19 kD species comigrates with N, and its formation depends upon the presence of microsomes, consistent with the proposal that N derives from F by signal cleavage and a further internal cleavage. The overall pathway for formation of the predominant forms of *hh* protein observed *in vivo* thus appears to involve signal cleavage of F to generate U. U is then cleaved internally to form N and C, which are the predominant forms found *in vivo*. Further processing of the 25 kD C species might then generate the 16 kD C^* species, but whether this processing is a single cleavage event or not is not clear since Ab2 does not recognize the smaller 9 kD fragment that would result. The processing of C to generate C^* appears to occur with greater efficiency in imaginal discs as compared to embryos (compare lanes 15 and 18); this may be caused by the more extended mass isolation procedure of imaginal discs (O. M. Eugene, *et al.*, *Tissue Culture Assn. Man.*, 5: 1055, 1979).

EXAMPLE 2

AUTO-PROTEOLYSIS OF THE HEDGEHOG PROTEIN

The comigration of endogenous and *in vitro*-generated *hh* protein species suggested that *in vitro* processing is similar to that observed *in vivo*. FIGURE 2 shows limited sequence similarity between *hh* proteins and serine proteinases. *hh* protein sequences are aligned to residues 323 to 329 of the *D. melanogaster* protein and numbered as positions 1 to 7 (group A). Conserved *hh* residues are in bold letters. The catalytic histidines (A. J. Barrett, in *Proteinase inhibitors* A. J. Barrett, G. Salvesen, Eds. (Elsevier, Amsterdam, 1986) pp. 3-22) of mammalian serine proteinases (group B) are aligned to the invariant histidine at position 7 in *Hh* proteins. Abbreviations are as follows: C-S*hh*, chicken *Sonic hh* (R. D. Riddle, *et al.*, *Cell* 75: 1401, 1993); M-S*hh*, mouse *Sonic hh* (Y. Echelard *et al.*, *Cell* 75: 1417, 1993) (identical to *Hh*g-1; R v*hh*-1, rat v*hh*-1 (H. Roelink et al., *Cell* 76: 761, 1994); Z-S*hh*, zebrafish *Sonic hh* (S. Krauss, *et al.*, *Cell* 75: 1431, 1993) (identical to *shh*) and zebrafish v*hh*-1, (H. Roelink et al., *supra*); *twhh*, no other abbreviation; M-D*hh*, mouse *Desert hh* (Y. Echelard et al., *Cell* 75: 1417, 1993); M-I*hh*, mouse *Indian hh* (Y. Echelard et al., *supra*); CHT, bovine chymotrypsin; TRP, bovine trypsin; ELA, porcine elastase; UKH, human urokinase; C1R, human complement factor 1R; C1S, human complement factor 1S; MCP, rat mast cell protease; FAX, human blood clotting factor X; TPA, human tissue plasminogen activator.

Figure 2 shows that a seven residue region of *hh* coding sequence (residues 323 to 329 in the *Drosophila* protein) displays some similarity to the sequences of serine proteases. This region lies approximately two thirds of the distance from the signal cleavage site to the carboxy-terminus, and includes Thr and His, residues (positions 4 and 7 in FIGURE 2) that are invariant among all *hh* sequences from all species. In the serine proteases, this conserved sequence contains an invariant His that acts as a general base in catalysis (A. J. Barrett, in *Proteinase inhibitors* A. J. Barrett, G. Salvesen, Eds. Elsevier, Amsterdam, 1986, pp. 3-22).

To determine whether this invariant His residue in the *hh* protein indeed plays a role in auto-proteolysis, two proteins from *E. coli* were purified: one carried the wild type sequence and the other a substitution of an Ala codon for the His codon at position 329 (H329A). Both of these proteins were engineered to contain a hexa-histidine tag at the amino terminus fused to *Drosophila* sequences extending from a residue just before the signal cleavage site to the carboxy-terminus (residues 83 to 471; the wild type form of this protein is referred to as His_6-U). Both proteins were extensively purified under denaturing conditions using a Ni^{++} -chelating matrix. FIGURE 3(A) is a coomasie blue stained polyacrylamide gel that shows the production and purification of His_6-U and His_6-U_{H329A} proteins from *E. coli*. Samples were molecular weight markers (lanes 1 and 2); lysates of *E. coli* cells carrying the His_6-U expression construct without (lane 3) and with (lane 4) induction by IPTG; purified His_6-U protein (lane 5); lysates of *E. coli* cells that carry the His_6-U_{H329A} expression construct without (lane 6) and with (lane 7) induction by IPTG; purified His_6-U_{H329A} protein (lane 8). Purified proteins were essentially homogeneous except for several minor species of lower relative mass; these species are endogenous breakdown products of the full-length proteins since they were absent in uninduced extracts and were detectable with *hh* antibodies. FIGURE 3 (B) is an immunoblot detected with Ab2 showing transfected S2 cells induced to express *hh* (lane 1); His_6-U and His_6-U_{H329A} proteins incubated in cleavage reaction buffer for 0 hours (lanes 2 and 5), for 20 hours (lanes 3 and 6), and for 20 hours in the presence of 20 mM TAME (a serine protease inhibitor) (lanes 4 and 7). Upon incubation the His_6-U, but not the His_6-U_{H329A} protein, released a fragment presumed to be C on the basis of reactivity with Ab2 and co-migration with C produced in S2 cells. Release of C (lane 3) was only partially inhibited by TAME.

Preliminary proteinase inhibitor studies have been performed on *in vitro* translated *Hh* protein by adding various inhibitors at the start of the translation reaction. These studies have been complicated by the fact that numerous protease inhibitors lower or block translation efficiency. In some cases the effectiveness of an inhibitor was assayed by determining if addition of an inhibitor to a completed translation reaction will inhibit the

self-processing that normally continues to occur. At this time we can only state the following with certainty: (i) the serine protease inhibitor TAME (p-toluenesulfonyl-L-arginine methyl ester) inhibits auto-proteolysis of in-vitro translated *Hh* protein; (ii) soybean trypsin inhibitor, a, anti-trypsin, aprotitin, leupeptin, and E-64 do not block auto-proteolysis of translated *Hh* protein; and (iii) TAME partially inhibits auto-proteolysis of purified His_6-U protein (FIGURE 3, panel B).

As seen in FIGURE 3B, upon dilution of denaturant the wild type protein but not the H329A mutant protein released a 25 kD species detectable by Ab2 and identical in mobility with the C species produced from *in vitro* translations and various *in vivo* sources. This cleavage was also observed when the wild type protein was purified and renatured by other protocols and cleaved under distinct conditions. Plasmids encoding the His_6-U and $His6\text{-}U_{H329A}$ proteins were generated by inserting sequences corresponding to residues 83 to 471 from the wild-type or *hh* H329A ORF into the pRSETB expression vector (Invitrogen). Proteins were induced in BL21(DE3)/pLysS *E. coli* cells as described (F. M. Ausubel et al., *supra).* The basic purification was performed on Ni-NTA agarose beads (Qiagen) by a denaturing protocol with the use of 6 M guanidinium HCl and 8 M urea essentially as recommended (a detailed protocol of exact conditions used is available upon request). Washes contained 0.2 percent Tween 20 and 5 mM b-mercaptoethanol. The final wash buffer was: 6 M urea, 100 mM Tris, 500 mM NaCl, 20 percent glycerol, (pH 7.4). Elutions were with the final wash buffer containing 250 mM imidazole. In vitro cleavage reactions were performed by incubating the purified protein (diluted 1:30 in the final mix) in cleavage buffer [50 mM Tris, 500 mM NaCl, 5 percent glycerol, 0.2% Triton X-100, 50 mM DTT, (pH 7.4)]. To isolate soluble full-length His_6-U protein free from denaturants or detergents, additional steps were taken (this refers to the other renaturation protocols mentioned in the text). Full-length protein from the eluate described above was further purified from breakdown products by precipitation, by urea removal through dialysis. The precipitate was then re-solubilized in a buffer containing guanidinium HCl and loaded onto another Ni-NTA agarose column. After washing as described, the protein was re-folded (while attached to the beads) by gradual dilution of urea (from 6M to 0.5M)

with dilution buffer [(100 mM Tris, 500 mM NaCl, 20 percent glycerol, (pH 7.4)] over an 8 hour period at 4° C. The protein was eluted with dilution buffer containing 250 mM imidazole and 0.5M urea. The eluate was dialyzed in 100 mM Tris, 150 mM NaCl, 10 percent glycerol, (pH 7.4) at 4° C and stored at -70° C.

EXAMPLE 3

MAPPING THE AUTO-PROTEOLYTIC FUNCTIONS OF *hh*

To more precisely define the domain of the *hh* protein responsible for this auto-proteolytic event, the effects of several distinct types of mutations upon *in vitro* processing were examined. The most informative mutation was a deletion that removes residues 89 to 254 (Δ89-254), which together constitute most of the amino acids within the portion of the molecule presumed to form the N fragment. I*n vitro* translations of wild-type and mutant *Hh* proteins from *Drosophila* (FIGURES 4 A-C) and zebrafish (FIGURE 4D) are shown. The locations of mutations and cleavage sites (arrows) in these proteins are schematically illustrated (FIGURE 4E). In the *Drosophila* protein (FIGURES 4A, B, and C), auto-proteolysis is blocked or severely inhibited by several mutations in the COOH-terminus (H329A, 294 trunc, 410 trunc, flu408 and 456 trunc), but is unaffected by a large deletion (Δ89-254) or insertion of a flu-tag epitope trimer (flu227) in the NH_2-terminus. Auto-proteolysis thus depends primarily on residues within the C fragment (sequences to the right of the cleavage site in the diagram below; see FIGURE 1). Furthermore, the H329A/flu227 double mutant is not cleaved by wild-type protein in a mixing experiment (lane 11), suggesting an intramolecular mechanism for auto--proteolysis. *Hh* proteins encoded by the zebrafish genes *twhh* and *shh* display a pattern of processing (D) similar to that of the *Drosophila* protein although the NH_2-terminal fragment of each zebrafish protein (23 kD for *twhh* and 22 kD for *shh*) has a lower apparent mass than the COOH-terminal fragment (25 kD for *twhh* and *shh*). This is the result of a shorter stretch of residues that precedes the signal sequences as compared to the *Drosophila* protein. Processing is blocked by H273A and H270A mutations in *twhh*

and *shh* proteins respectively (analogous to the H329A mutation in the *Drosophila* protein), which suggests an auto-proteolytic processing mechanism is used similar to that observed for the *Drosophila* protein.

In vitro translations were performed with the use of the TNT coupled transcription-translation system (Promega). ^{35}S methionine (DuPont NEN) was used for detection by autoradiography. In the heparin binding experiment (FIGURE 8C), *in vitro* translation lysate with microsomes that produce wild-type *Hh* protein was added to heparin agarose (Sigma) or Sepharose CL-4B (Pharmacia) beads pre-equilibrated with heparin binding buffer (HBB; 20 mM Tris (7.4), 150 mM NaCl, 0.1 percent Triton X-100). Samples were incubated at 4° C for four hours with gentle rocking. After pelleting the beads, supernatants in some samples were analyzed (lanes 2 and 4). The beads were then washed 5 times with chilled HBB and samples (lanes 3 and 5) were subsequently eluted at 80° C for 10 minutes in SDS PAGE loading buffer (F. M. Ausubel et al., *supra)*.

All mutations in the *hh* gene were generated in the plasmid pF1 (J. J. Lee, *et al., supra)*. Mutations in the zebrafish *twhh* and *shh* genes were generated with the original cDNA clones as described (Ekker, *et al., Current Biology*, 5(8): 944,1995). All point mutations were generated with the use of recombinant circle PCR (D. H. Jones and S. C. Winistorfer, *Biotechniques* 12: 528, 1992). The flu408 and flu227 mutations were generated by inserting a trimer of the influenza hemagglutinin antigen (42 residues for flu408 and 43 residues for flu227) into the AlwN I and Bgl I sites present in the *hh* ORF (nucleotide positions 1604 and 1058 respectively) (J. J. Lee, *et al., supra*). The Δ89-254 mutation was generated by removing sequences between the EcoN I site (644) and the Pml I site (1145). The 294 trunc mutation was generated by removing sequences between the Acc I site (1265) and the Xcm I site (1792). The 410 trunc mutation was previously generated and identified as Hh_{410} (J. J. Lee, *et al., supra*). To map the mutation in the *hh*^{13E} allele (base change C_{1756} to A; coding change Tyr_{157} to STOP), DNA isolated from *hh*^{13E}/TM3 was used to seed PCR reactions generating regions of the *hh* ORF and flanking

sequences, which were subcloned into Bluescript KSM (Stratagene). Six clones each, derived from two different PCR amplifications were sequenced.

As seen in lanes 1 and 2 of FIGURE 4A, this construct generates a full length species of a mobility corresponding to the expected relative mass of 33 kD, and two cleaved products whose apparent relative masses (25 and 9 kD) sum to give the relative mass of the larger species. The smaller of the cleaved products will occasionally migrate as two bands as seen in Fig 4A. We have chosen the lower of the two bands between the 14.3-kD and 6.2-kD markers for our molecular weight measurement. The larger of the two cleaved products comigrates with the C species produced from the wild type protein, suggesting that the Δ89-254 *hh* protein contains the residues normally present in C and all of the determinants required for auto-proteolysis, including the normal cleavage site; most of the residues within N are dispensable for auto-proteolytic activity.

In contrast, lesions affecting residues presumed to lie within C block auto-proteolysis *in vitro*. All mutations tested by *in vitro* translation were also examined in S2 cells by immunoblotting. In all cases the patterns of cleavage in S2 cells were identical to those observed in translations except that C* was always present whenever C was formed. The former fragment was not observed in translations. These include the H329A mutation described above, a mutation that inserts an influenza virus epitope between residues 408 and 409 (flu408), and three mutations that cause premature termination of the protein at the carboxy terminus. The two most severe truncations, 294 trunc and 410 trunc, are mutations generated *in vitro*. They cause a loss of 177 and 61 residues, respectively, from the carboxyl-terminus of the protein, and neither undergoes proteolysis. The 456 trunc *hh* protein is like that encoded by the EMS-induced hh^{13E} mutant allele, which results in the loss of 15 residues from the carboxy-terminus of the protein. This protein undergoes auto-proteolysis, as demonstrated by the appearance of a 24 kD band in place of C, but the efficiency of the reaction is much impaired *in vitro* (FIGURE 4B). Auto-proteolysis of the *hh* protein relies mainly upon residues within C; deletion or alteration of residues

within this domain is associated with reduced efficiency of processing, and one such deletion appears to be the cause of the *hh*^{13E} mutation.

The sequence homology and auto-proteolytic function of the full length *hh* protein suggested the possibility that F or the C fragment is a sequence-specific protease. As a first step in clarifying the mechanism of auto-proteolysis, an influenza virus epitope tag was introduced into the N-terminus of a *hh* open reading frame that also carried a H329A mutation. FIGURE 4C shows that the insertion of the epitope tag alone does not interfere with auto-proteolysis (lane 9), and yields a normal C fragment and an N fragment of increased relative mass (compare to wild type in lane 12). The protein carrying both mutations does not undergo proteolysis (lane 10), and since the epitope-tagged N fragment migrates differently from N, this double mutant provides an ideal substrate to look for intermolecular cleavage upon mixture with a wild type sequence. Lane 11 shows that in such a mixture, although normal N is formed, no tagged N can be detected. Thus, in this experiment, no appreciable intermolecular cleavage occurs. We also failed to detect intermolecular cleavage in the following two experiments: (i) co-transfection of wild type and 410 trunc sequences into S2 cells (the cleaved 410 trunc protein would yield a smaller and therefore identifiable form of C); (ii) mixing of excess unlabelled, purified His_6-U protein with labelled, *in vitro* translated H329A mutant protein. Thus, although an intermolecular mechanism for regulation of auto-proteolysis or for cleavage of other proteins can not be ruled out, the current evidence suggests that cleavage of the *hh* protein occurs predominantly by an intramolecular mechanism.

The *hh* gene has been broadly conserved in evolution, with single homologues unidentified in a wide variety of invertebrate species and multiple distinct homologues in each of several vertebrate species (Y. Echelard et al., *Cell* 75: 1417, 1993; S. Krauss, *et al.*, *Cell* 75: 1431, 1993; H. Roelink et al., *Cell, supra*). As seen in FIGURE 2, all of these coding sequences contain an invariant histidine and other conserved residues at a position corresponding to H329 in the *Drosophila* protein. In addition, the protein encoded by at least one of the mouse genes appears to be processed *in vivo* to yield two

smaller species in a manner resembling the *in vivo* processing of the *Drosophila* protein. To determine whether auto-proteolysis may also play a role in vertebrates we examined the behavior of proteins encoded by two distinct *hh* homologues from the zebrafish, tw*hh* and s*hh*. FIGURE 4D demonstrates that when these sequences are translated *in vitro*, smaller species are generated whose relative masses sum to yield approximately the relative mass of the full length protein (lanes 1 and 3). As seen in lanes 2 and 4, this cleavage reaction is blocked by substitution of Ala codons for the His codons at positions corresponding to H329 in *Drosophila* (see FIGURE 2). Vertebrate *hh* proteins thus appear to be processed by a similar mechanism as the *Drosophila* protein.

EXAMPLE 4

ROLE OF AUTO-PROTEOLYSIS IN EMBRYOS

Numerous functions for the *hh* gene have been described in *Drosophila*. At the morphological level these include a role in patterning of larval cuticular structures and adult structures such as the eye and appendages (C. Nüsslein-Volhard and E. Wieschaus, *Nature* 287: 795, 1980; and J. Mohler, *Genetics* 120: 1061, 1988).; the mechanistic basis for thee morphological effects involves signaling for maintenance or induction of gene expression in embryos and imaginal discs (J. J. Lee, *supra*; T. Tabata and T. B. Kornberg, *Cell* 76: 89, 1994; and K. Basler and G. Struhl, *Nature* 368: 208, 1994). To ascertain the importance of auto-proteolysis for these functions, the H329A mutant gene under control of the hsp 70 promoter was introduced by P element-mediated transformation into the *Drosophila* germline. The hs*hh* H329A construct was made identically to the hs*hh* construct with the use of a *hh* ORF fragment containing the H329A mutation. Transgenic flies were generated from a y^1 w^{1118} parental strain using standard methods of P element mediated transformation (A. C. Spradling and G. M. Rubin, *Science* 218: 341 1982). A line, HA3, carrying the hs*hh* H329A P element on the second chromosome was maintained as a homozygous stock. T assay for expansion of *wg* stripes, embryos collected at 4 to 6 hours after egg laying (AEL) at 25° C were subjected to the following

heat shock protocols prior to fixation. Embryos receiving single shocks (10 or 30 minutes at 37° C) were allowed to recover for 1 hour at 25° C. Embryos receiving double shocks (two 10 minute or two 30 minute shocks at 37° C) were allowed to recover 90 minutes after the first shock and 40 minutes after the second (Both recoveries were at 25° C. The double 30 minute protocol was as previously described, (S. Krauss, *supra*). *In situ* hybridizations were performed as described (D. Tautz, *Chromosoma* 98: 81, 1989) using a *wg* specific probe (D. T. Chang et al., *supra*). Embryos assayed for cuticle phenotype were heat shocked 6 to 8 hours AEL for 30 minutes at 37° C, allowed to develop at 25° C for 36 hours and then processed and mounted as described (M. Ashburner, *Drosophila: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, New York, 1989). Immunolocalizations (single or double stains) were performed as described. With the use of affinity purified Ab1 or Ab2 for the primary antibody and alkaline phosphatase (AP) or horseradish peroxidase (HRP) conjugated anti rabbit or mouse IgG (Jackson Immunoresearch) for the secondary. Embryos from a hh^{13E}/TM3 *ftz*-lacZ (the balancer chromosome was from the Bloomington Stock Center, strain 3218) stock homozygous for the hh^{13E} allele were identified by the lack of staining with an anti b-galactosidase antibody (Promega) in a double stain with Ab2 (FIGURE 9, panel D). Staining in FIGURE 9, panels B and C were performed formaldehyde fixed Canton-S embryos with the use of an AP conjugated anti-rabbit IgG secondary. Although standard formaldehyde fixation was generally used, heat and acid-formaldehyde fixation also gave similar results. GST fusion proteins containing either residues 83 to 160 or 300 to 391 from the *Hh* protein were expressed in *E. coli*, purified as recommended (F. M. Ausubel et al., *supra*), and used to immunize rabbits by standard methods. The antibodies were affinity purified on a column of His_6-U protein (Harlow and Lane, *supra*) linked to Affi-Gel 10 beads (Bio-Rad). The purification was performed as described (Harlow and Lane, *supra*) except that the acid and base elutions contained 10 percent dioxane. Biotinylated *hh* antibodies were prepared by purifying the rabbit antisera over a protein A column, followed by biotinylation with the use of the Immunoprobe biotinylation kit (Sigma). Immunoprecipitations were performed as described (Harlow and Lane, *supra*) with the use of cold RIPA lysis buffer containing 0.25 mM PMSF and 5 mM EDTA for tissue

homogenization. Lysates were precleared twice with pre-immune rabbit serum plus protein A beads (Gibco BRL). Affinity purified antibodies or pre-immune serum was then added, and the immunoprecipitation was performed with protein A beads, with the use of NP-40 lysis buffer for the washes.

FIGURE 5 (A) and (B) are immunoblots developed with the use of Ab1 and Ab2 antibodies respectively. Lanes 1 and 6, induced untransfected S2 cells; lanes 2 and 7, transfected S2 cells induced to express *hh*; lanes 3 and 8, heat shocked wild-type embryos; lanes 4 and 9, heat shocked hs*hh* embryos; lanes 5 and 10, heat shocked hs*hh* H329A embryos. In heat shocked hs*hh* embryos, the wild-type *Hh* protein is both induced and properly processed to generate the U, N, C and C* species seen in other expression contexts. In contrast, the H329A is induced but not appreciably processed in hs*hh* H329A embryos (the low levels of processed species in lanes 5 and 10 are probably from endogenous *hh* expression since they are seen at identical levels in heat shocked wild-type embryos in lanes 3 and 8).

FIGURE 5 shows that heat shock induction results in the formation of an abundant species that corresponds to U based on its mobility and its interaction with Ab1 and Ab2 (lanes 5 and 10). In contrast, induction of wild type *hh* protein using a similar contruct resulted in similar levels of the N and C processed products (lanes 4 and 9), with very little uncleaved U. Thus, as observed *in vitro* and in S2 cells, the H329A mutation in embryos appears to greatly reduce the efficiency of auto-proteolytic cleavage of the *hh* protein.

In FIGURE 6, the embryonic distribution of *wingless* (*wg*) RNA as revealed by in situ hybridization is shown in FIGURE 6 (A) wild-type (homozygous y^1 w^{1118}), (B) hs*hh*, and (C) hs*hh* H329A embryos that were exposed to two 10 minute heat shocks separated by a 90-minute recovery period. Wild-type embryos showed little change in *wg* expression, whereas the wild-type protein and, to a lesser extent, the H329A protein each induced ectopic *wg* expression (Table 1). Panels (D), (E), and (F) show the dorsal surfaces of y^1

w^{1118}, hs*hh*, and hs*hh* H329A larvae, respectively, at the level of the fourth abdominal segment. These larvae were shocked for 30 minutes as embryos and allowed to complete embryogenesis. Cuticle cell types (1°, 2°, 3°, and 4°) are labeled as described (J. Heemskerk and S. DiNardo, *Cell* 76: 449, 1994). Note the expansion of 2° cell types (naked cuticle) at the expense of 3° and some 4° types in the hs*hh* embryo (E) under conditions where the phenotype of hs*hh* H329A embryos (F) is identical to that of control embryos (D).

Perhaps the earliest known requirement for *Hh* protein is in maintenance of an adjacent stripe of *wingless* (*wg*) gene expression in each embryonic segment (A. Martinez Arias, *et al.*, *Development* 103: 157, 1988; and S. DiNardo, *et al.*, *Nature* 332: 604, 1988). This requirement is deduced from the loss of *wg* expression when *hh* function is absent; in addition, the ubiquitous expression of wild-type *Hh* protein induces expansion of the domain of *wg* gene expression (P. W. Ingham, *Nature* 366: 560, 1993). The effects of the H329A mutation upon *wg* expansion were examined by heat shocking embryos carrying the H329A mutant construct in parallel with embryos containing the wild-type construct. Although the H329A mutant protein is able to induce some expansion of the *wg* domain, the efficiency of this activity is impaired relative to that of the wild-type protein (FIGURE 6, B and C; Table 1). The difference in efficiency ranges nearly as high as threefold depending upon the heat shock regime, and these results suggest that a uto-proteolysis of the *Hh* protein is important for optimal activity in embryonic signaling to induce *wg* expression.

TABLE 1

Wild-type and mutant *hh* activity in embryonic induction of *wg* expression*

	minutes of heat shock			
	10	30	10/10	30/30
hs*hh*	1.0 ± 0.3 (93)	1.5 ± 0.6 (120)	2.9 ± 0.3 (41)	2.8 ± 0.4 (54)
hs*hh* H329A	0.7 ± 0.5 (190)	0.9 ± 0.4 (111)	1.1 ± 0.4 (145)	1.9 ± 0.5 (93)

* Expansion of *wg* expression beyond wild-type controls is given as average number of cell diameters ± standard deviation with number of embryos scored in parentheses.

The effects of *Hh* protein on the patterning of cuticular structures are most clearly visible on the dorsal surface of the larva, where four distinctive cell types can be identified in each parasegment. These cell types have been designated 1°, 2,° 3°, and 4°, from anterior to posterior, with *hh* transcription occurring in precursors of the 1° cells (J. Heemskerk and S. DiNardo, *supra*). Differentiation of the first three cell types was shown to be dependent upon *hh* gene function, and it has been proposed that the fates of these cells are determined by the concentration of *Hh* protein, with highest concentrations producing the 1° fate, intermediate concentrations producing the 2° fate, and the lowest concentrations producing the 3° fate (J. Heemskerk and S. DiNardo, *supra*). This proposal was supported by observations that the most anterior cell types display the greatest sensitivity to a reduction of *hh* expression, and that all of the 3° and some of the 4° bristles are replaced by naked cuticle characteristic of the more anterior 2° cell type when *hh* is expressed ubiquitously at high levels. We have reproduced suppression 3° and some 4° fates by heat shock induction of embryos that carry our wild-type construct (FIGURE 6E), but find that the H329A mutant is unable to alter cell fates in the dorsal cuticle of the larva (FIGURE 6F). Auto-proteolysis, or perhaps some other function blocked by the H329A mutation, thus appears to be essential for the patterning influence of *Hh* protein upon the dorsal cuticle.

EXAMPLE 5

EFFECTS OF THE H329A MUTATION UPON SIGNALING IN IMAGINAL DISCS

Studies of H329A mutant protein were extended to the function to the patterning of adult structures and signaling within imaginal discs. In the eye imaginal disc *hh* function is required for appropriate development of pattern (J. Mohler, *Genetics* 120: 1061, 1988; J. J. Lee, *supra*; and J. Mohler and K. Vani, *supra*) and more recently has been shown to control progression of a wave of differentiation via induction of *decapentaplegic* (*dpp*)

gene expression in the morphogenetic furrow of the eye (U. Heberlein, *et al.*, *Cell* 75: 913, 1993; and C. Ma, *et al.*, *Cell* 75: 927, 1993). In leg and wing discs, ectopic expression of *hh* has also been shown to yield pattern duplications and defects and is associated with induction of ectopic expression of other signaling molecules normally expressed in a zone along the anterior/posterior compartment boundary (T. Tabata and T. B. Kornberg, *Cell* 76: 89, 1994; and K. Basler and G. Struhl, *Nature* 368: 208, 1994).

For studies of signaling in imaginal discs, a thermal cycler was utilized to subject larvae carrying heat shock-inducible *hh* constructs to successive rounds of heat shock and recovery. The effects of temperature cycling upon expression of *dpp* and *wg* in imaginal discs was examined by monitoring ß-galactosidase expression from a reporter gene carrying *dpp* promoter sequences or from an enhancer detector P element inserted in the *wg* gene. In FIGURE 7, X-gal staining was used to follow expression of *wg* FIGURE 7 (A-C) or *dpp* FIGURE 7 (D-L) in imaginal discs of late third-instar larvae that carry *wg-lacZ* or *dpp-lacZ* reporter genes. Leg (A-F), wing (G-I) and eye-antennal discs (J-L) from control larvae (A, D, G, J), larvae carrying the hs*hh* transgene (B, E, H, K) and larvae carrying the hs*hh* H329A transgene (C, F, I, L) are displayed. In all panels anterior is to the left. Arrows highlight the following features: an ectopic patch of *dpp* expression in the anterior compartment of wing discs in hs*hh* H329A larvae (I); and an ectopic band of *dpp* expression in eye portion of the eye-antennal disc anterior to the morphogenetic furrow (marked by the other band of *dpp* expression more posteriorly) in hs*hh* larvae (K). Expansion into the anterior compartment of *wg* expression in leg discs, and *dpp* expression in leg and wing discs in hs*hh* larvae is similar to that described for the ectopic expression of *hh*. Morphological changes in the anterior compartment of leg (B and E) and wing discs (H) were also as described (K. Basler and G. Struhl, *supra*). In contrast, discs from hs*hh* H329A and control larvae showed very little change in *wg* and *dpp* expression, even under prolonged heat shock conditions and morphological changes were never observed. (M-O) The eye phenotypes of adult control (M), hs*hh* (N) and hs*hh* H329A (O) flies that were shocked during larval development in a manner similar to that of the imaginal disc experiments above. Duplicated eye structures wer observed in hs*hh*

flies, but never in hs*hh* H329A flies. The arrow in (N) points to a thin strip of cuticle between the two eye structures. Other deformities were also seen in hs*hh* flies (for example, compare the thorax in N to M).

Virgin female flies from the homozygous lines hs*hh* (D. T. Chang et al., *Development*, 1994, in press), hs*hh* H329A, and y^1 w^{1118} were crossed to males from the homozygous BS3.0 line (bearing a P element *dpp* reporter construct on the 2nd chromosome, referred to as *dpp-lacZ*) (R. K. Blackman, *et al.*, *Development* 111: 657, 1991) or the line *y*; *Sco/CyO, enlacZ11::wg* (bearing a *wg* reporter P element enhancer trap on a second chromosome balancer; called *wg-lacZ*) (J. A. Kassis, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 89: 1919, 1992). Progeny were grown at 25°C in aerated 0.5-ml microcentrifuge tubes containing yeast paste until the late second instar or early third instar stage of larval development. The larvae were then cycled continuously at 37° C for 30 minutes followed by 25° C for 90 minutes in a Perkin-Elmer thermal cycler until they reached the late third instar stage. They were subsequently dissected and stained with X-gal as described (M. Ashburner, *supra*) or allowed to grow to adulthood for phenotypic analysis.

As shown in FIGURE 7A, *wg* expression normally occurs in a ventral sector of the leg disc along the anterior/posterior compartment boundary while *dpp* is expressed in the dorsal portion of the disc along this boundary (FIGURE 7D). Although thermal cycling of larvae carrying the wild-type *hh* gene produced abnormal leg disc morphology and extensive ectopic expression of both target genes, as previously reported for ectopic *hh* expression (FIGURE 7B and E), the H329A construct produced little if any detectable difference in these patterns of expression (FIGURE 7, C and F). Ectopic *hh* expression in the wing disc also leads to morphological changes and expanded expression of *dpp* (compare FIGURE 7, G and H), but the H329A construct produced only an occasional small patch of anterior ectopic expression (FIGURE 7I).

Ubiquitous expression of wild-type *hh* also leads to ectopic expression of *dpp* in the eye-antennal disc (compare FIGURE 7, J and K). In the antennal portion of this disc the

expansion of *dpp* expression resembles that observed in leg discs. In the eye portion of the disc *dpp* expression is observed at its normal location in the furrow; however, ectopic expression also occurs in the form of a second dorso-ventral band at a location somewhat anterior to the furrow, thus giving the appearance of an eye disc with two morphogenetic furrows (FIGURE 7K). Indeed, in adults derived from temperature-cycled larvae that carry the wild-type *hh* construct, an apparently duplicated eye structure such as that in FIGURE 7N can be observed, with two eye structures separated by a thin strip of cuticle (arrow). The H329A mutant protein, in contrast, did not induce expansion of *dpp* expression in either portion of the eye-antennal disc (FIGURE 7L), and does not induce eye duplications or cuticle defects in the adult (FIGURE 7O).

The experiments described thus far comprise multiple series of larvae subjected to two days of thermal cycling followed by immediate dissection for analysis of imaginal structures or further incubation at constant temperature for analysis of adult structures. Although the H329A protein appeared to have little activity in these experiments, the small patch of ectopic *dpp* expression induced in the wing disc (FIGURE 7I, arrow) suggested that some residual activity remained. This suggestion was borne out in a similar experiment involving three days of cycling prior to dissection: the H329A protein clearly displayed some *dpp*-inducing activity in this experiment, presumably as a result of the higher amounts of protein that accumulated during the longer cycling period. The wing in particular, but also other imaginal discs, displayed low and variable amounts of ectopic *dpp* expression. This expression in all cases was far less extensive than that observed for the wild-type construct examined in parallel; furthermore, morphological deformations of the imaginal discs, although quite common with the wild-type protein, were extremely rare with the H329A protein. Although its potency is greatly reduced relative to wild-type, the H329A protein retained at least some activity in early embryonic and imaginal disc induction of *wg* and *dpp* expression; in contrast, even under heat shock conditions far more severe than those required for effects by the wild-type protein, the H329A mutant remained completely inert with respect to the re-specification of cell fates in the dorsal cuticle of the larva.

EXAMPLE 6
DIFFERENTIAL RELEASE OF N AND C INTO CULTURED CELL SUPERNATANTS

A puzzling feature of *hh* function is its apparent short-range action in settings such as embryonic and imaginal disc signaling to *wg* and *dpp*, and longer-range action in other settings, such as patterning of the dorsal larval cuticle. These observations and the existence of two major protein products in vivo prompted us to look for differences in the solubility or diffusibility of N and C expressed in S2 cultured cells. FIGURES 8 (A) and (B) are immunoblots of cell pellets (lane 1) or supernatants (lane 2) from transfected S2 cell cultures expressing *Hh* protein, developed with Ab1 (A) and Ab2 (B). Samples in each lane were from the same volume of resuspended total culture. Whereas N remained mostly associated with the cell pellet (compare lanes 1 and 2 in A), C was nearly quantitatively released into the supernatant (compare lanes 1and 2 in B). U displayed partitioning properties in between those of N and C (A and B). (8C) demonstrates the heparin binding activity of various *Hh* protein species generated by *in vitro* translations with microsomes. Samples were: total translation mix (lane 1); supernatant after incubation with heparin agarose or agarose (control) beads (lanes 2 and 4); and material eluted from heparin agarose or agarose beads after washing (lanes 3 and 5). F, U, N_{ss} and N fragments are depleted from reactions incubated with heparin agarose but not agarose beads (compare lanes 2 and 4 to 1), and the same species subsequently can be eluted from the heparin agarose but not the agarose beads (compare lanes 3 and 5 with lane 1). FIGURES 8, A and B indeed show that these proteins behave differently, with most of the N fragment remaining cell-associated and all, or nearly all, of C being released into the culture supernatant.

One possible explanation for this differential behavior might be association of the N fragment with extracellular matrix proteins on the surfaces of the S2 cells. Accordingly, the relative affinity f these two proteins for heparin agarose was examined, since heparin binding is a common property of proteins that associate with the extracellular matrix.

Given the obvious difficulty in obtaining soluble N from cultured cells, in vitro translation in the presence of microsomes was used to generate soluble, labelled N and C. As shown in FIGURE 8C, N but not C is depleted from these translation extracts by treatment with heparin agarose beads, while treatment with unmodified agarose beads did not deplete either fragment. Furthermore, N but not C was retained upon the heparin agarose beads upon extensive washing with a solution that contains 0.1% Triton X-100 and 150 mM NaCl; in contrast, neither fragment was retained by unmodified agarose. N, but not C, binds tightly to heparin, and this behavior suggests that the low concentration of N released into culture supernatants may be the result of binding to the extracellular matrix. Another mechanism that might contribute to the differential release of N and C into culture supernatant would be the expression in S2 cells of a receptor for N but not for C. Our current data can not distinguish these possibilities.

EXAMPLE 7

DISTINCT EMBRYONIC LOCALIZATIONS OF N AND C

The differential release of N and C into cultured cell supernatants suggested the possibility that these fragments might also be differentially localized in embryos. Previously reported *hh* protein localizations utilized either antibodies specific for N epitopes or antibodies unable to distinguish between N and C. FIGURE 9 shows the differential localizations of N and C in embryos by in situ localization of the *hh* transcript. FIGURE 9 (A) is shown in comparison to the distribution of N and C epitopes detected with Ab1 and Ab2 in panels (9B) and (9C), respectively. Note that the distribution of N and C epitopes span approximately one-third and one-half of each segmental unit respectively, while the transcript is limited to approximately one-quarter of each unit. In (9D), the localization of C epitopes in embryos homozygous for the hh^{13E} allele is detected with the use of Ab2. C epitopes in this mutant, which displays impaired auto-proteolytic activity are more restricted, and resemble the wild-type localization of N. Homozygous hh^{13E} embryos were identified by loss of a marked balancer from a heterozygous parent stock. All embryos are at mid to late stage 9 (extended germ-band).

FIGURE 9B shows in accordance with these reports, Ab1, which is specific for N epitopes, reveals a segmentally localized distribution that is slightly broader than that of the *hh* transcript at the same stage (FIGURE 9A). Also consistent with these reports, we observed that N epitopes at later stages accumulate in large punctate structures. Our analysis here concentrates on the earlier stage, when antibody staining is weaker but before formation of the invaginations and grooves that later crease the epidermis and thereby complicate the interpretation. Ab2 was also utilized to detect C-specific epitopes with a variety of fixation and staining procedures. Although detection of C epitopes above background is more difficult than for N, we consistently observed a segmentally modulated pattern, albeit with a broader distribution than N (FIGURE 9C). This localization is also distinctive in that C epitopes at early or late stages are not found in the punctate structures characteristic of N.

The *hh*^{13E} mutation encodes a prematurely truncated protein that is missing 15 residues normally present at the COOH-terminus. Because this protein displays a much reduced efficiency in auto-proteolysis the distribution of C in this mutant background was examined. FIGURE 9D shows that C epitopes in a homozygous *hh*^{13E} embryo (identified by absence of a marked balancer) are distributed in a much tighter segmental pattern than in wild-type. This localization resembles that of N, and we thus conclude that the broad distribution of C epitopes normally seen is altered in *hh*^{13E} by retention of the uncleaved precursor near the site of synthesis.

EXAMPLE 8

THE ROLE OF AUTO-PROTEOLYSIS IN BIOGENESIS OF ACTIVE HEDGEHOG PROTEIN

In addition to signal cleavage, the *hh* protein undergoes auto-proteolysis at an internal site to generate the predominant protein species observed *in vivo*. All or most of the amino acid residues required for this auto-proteolysis function map to C, the carboxy-terminal product of this internal cleavage. In an effort to determine the importance of auto-proteolysis for function, we introduced a single residue mutation (H329A) that blocks auto-proteolysis of the *hh* protein in vitro and demonstrated that both processing and function of this protein is impaired in vivo. Since similar levels of induced protein were detected from a strain carrying the wild-type construct or from several strains carrying independent insertions of the mutant construct (FIGURE 5), the impaired function of the H329A protein relative to wild-type is not the result of reduced levels of expression. Further evidence in support of a role for auto-proteolysis derives from the effect of the *hh*^{13E} mutation, which reduces but does not eliminate auto-proteolysis of the *hh* protein *in vitro* (FIGURE 4). Correspondingly, the *hh*^{13E} mutation is associated with a phenotype of intermediate strength in vivo (J. Mohler, *supra*).

Curiously, the H329A *Hh* protein appears to retain weak activity in embryonic signaling to induce ectopic *wg* expression and, to a lesser degree, can function in imaginal disc

signaling for induction of ectopic *wg* and *dpp* expression. In contrast to its retention of at least some signaling functions in embryonic and imaginal tissues, the H329 protein is completely inert when assayed for the ability to reprogram cell fates in the dorsal cuticle of the larva.

The assays in which the H329A protein is active or partially active involve short-range signaling that normally occurs across one or at most several cell diameters; in contrast, the H329A protein fails to exert any effect upon patterning of the dorsal cuticle, a long-range activity that normally operates across most of the segment. Previous proposals to account for long-range patterning activities have suggested that *hh* expression induces other signaling molecules which are then responsible for executing the patterning functions (the signal relay model; see FIGURE 10A). FIGURE 10 shows a signal relay versus dual function models for *hh* protein action. In FIGURE 10 (A), the long-range effects of *hh* signaling are achieved indirectly through short-range induction of a second signaling molecule (X). Based on its biochemical properties and its restricted tissue localization, N is presumed to represent the active short-range signal while the role of C would be limited to supplying the catalytic machinery required for biogenesis of N. In (10B), the long- and short-range signaling functions of *hh* are supplied by the N and C proteins derived by internal auto-proteolysis of the U precursor. N is implicated in short-range signaling by retention near its cellular site of synthesis, while C is less restricted in its distribution and would execute long-range signaling functions. In both models, auto-proteolysis is required to generate fully active signaling proteins. See text for further discussion.

These proposals seek to maintain a consistent mode of *hedgehog* action by rationalizing the apparent long-range activities of *hh* products as indirect consequences of short-range signaling. Based on the distribution observed, the active molecule in this model might be N and the role of C would then be limited to supplying the catalytic machinery required for biogenesis of N.

Our evidence suggests an alternative model, the dual function model (FIGURE 10B), in which long- and short-range activities of the *hh* protein might be executed by N and C, the two predominant forms of the molecule observed *in vivo*. The nearly quantitative release of C fragment into the culture medium of *hh*-expressing S2 cells and its broad, though segmentally modulated distribution within embryos suggests that C might execute or contribute to long-range signaling functions. The N fragment, on the other hand, predominantly remains associated with the expressing S2 cells and also binds to heparin, which suggests a possible association with the extracellular matrix. These properties and the segmentally restricted embryonic distribution of N are suggestive of a role in the execution of short-range *hh* signaling activities. Since the vertebrate *Hh* proteins we tested also appear to be auto-processed and also carry predicted heparin binding sites just carboxy-terminal to their signal sequences (H. Roelink et al., *supra*), many aspects of the dual function model discussed here in the context of *Drosophila* development may also apply to *hh* protein function in vertebrate development.

Execution of short-range functions by N would be consistent with the observation that the H329A mutant protein has at least partial function in signaling for the induction of *wg* and *dpp*, since this mutation does not alter residues located in the amino-terminal portion of the protein that normally would give rise to N. The uncleaved H329A protein thus would carry all the residues that normally interact with a presumed receptor for N, although there might be some effect on the affinity of the interaction due to the presence of carboxy-terminal sequences, thus accounting for the decreased potency of the H329A protein. Alternatively, the partial function of H329A protein may derive from an extremely small fraction of protein that appears to be cleaved, a very faint band with identical mobility to C appears in *in vitro* translations with the H329A protein (FIGURE 4, lane 3). Execution of long-range functions by C is also consistent with our observations because long-range signaling might require the release of the C fragment or otherwise require the H329 residue for some function other than for cleavage.

When N is synthesized from a native construct (wild type hh), it remains primarily cell-associated (FIGURE 10C), however, N generated from a truncated construct in cultured cells predominantly enters the culture medium (FIGURE 10D) (For constructs, see Porter, *et al.*, *Nature*, 374:363, 1995). These results further confirm that autoprocessing by fragment C may regulate the degree of N association with the cell surface and therefore its range of action.

EXAMPLE 9
ISOLATION OF HEDGEHOG HOMOLOGUES

The mouse and human *hh*-like sequences were isolated by polymerase chain reaction (PCR) using primers degenerate for all possible coding combinations of the sequences underlined in FIGURE 1 of Chang, *et al.*, *(Development*, 120: 1994). PCR amplifications contained from 100 ng to 2 µg genomic DNA (depending upon the genome size of the species), 2 µM of each primer, 200 µM dNTPs (Pharmacia), 1X reaction buffer (Boehringer-Mannheim) and 2.5 units Taq polymerase (Boehringer-Mannheim) in 50 µl reactions. Amplification was as follows: 94°C 5 min, addition of Taq polymerase at 75°C, followed by 94°C 1 min, 52°C 1.5 min and 72°C 1 min for 30 cycles and a final extension of 72°C for 5 min. All PCR products were cloned into pBluescript (Stratagene) prior to sequence determination.

Mouse clones obtained in this manner contained 144 bases of sequence between the primer ends and were labelled with [α^{-32}P]dATP and used for high stringency screens of mouse cDNA libraries made from whole 8.5 dpc embryonic RNA and from 14.5 dpc embryonic brain in the λZAP vector (a gift from A. Lanahan). Several clones corresponding to *Hhg*-1 were isolated and the largest, 2629 bp in length (pDTC8.0), was chosen for sequence analysis using dideoxy chain termination (Sanger, *et al.*, 1977) and Sequenase v2.0 (US Biochemicals). Compressions were resolved by using 7-deaza-guanosine (US Biochemicals). Sequence analysis made use of the Geneworks 2.0 (IntelliGenetics) and MacVector 3.5 (IBI) software packages.

One of the three mouse clones, *Hhg*-1, when used as a probe, yielded a 2.0 kb clone from a 8.5 dpc mouse embryonic cDNA library and a 2.7 kb clone from a 14.5 dpc embryonic cDNA library. The 2.7 kb cDNA appears to represent a nearly full length mRNA because it corresponds to a 2.7 kb band detected by hybridization on a Northern blot. The largest methionine-initiated open reading frame within this cDNA encompasses 437 codons, and is preceded by one in frame upstream stop codon. Sequence comparisons indicate that the protein encoded by *Hhg*-1 is identical to the independently characterized mouse S*hh* (Echelard, *et al.*, *Cell*, 75:1417-1430, 1993) except for an arginine to lysine difference at residue 122. *Hhg*-1 also corresponds closely to the rat v*hh*-1 gene (97% amino acid identity; Roelink, *et al.*, *Cell*, 76:761-775, 1994), the chicken Sonic hedgehog (81% identity; Riddle, *et al.*, *Cell*, 75:1401-1416, 1993) and S*hh* from the zebrafish (68% identity; Krauss, *et al.*, *Cell*, 75:1431-1444, 1993; Roelink, *et al.*, *Cell*, 76:761-775, 1994). The PCR-generated fragments *Hhg*-2 and *Hhg*-3 appear to correspond to the Indian and Desert classes of mouse hedgehog genes, respectively (Echelard, *et al.*, *Cell*, 75:1417-1430, 1993).

Alignment of the *Hhg*-1 open reading frame with the two *Drosophila hh* sequences showed that all three proteins contain hydrophobic amino acid sequences near their amino-termini; the hydrophobic stretches within the *D. melanogaster* protein (residues 64 to 83) and within the mouse protein are known to act efficiently as signal sequences for cleavage (Lee, *et al*, *Cell*, 71:33-50, 1992). Both *Drosophila* signal sequences are unusual in their internal locations, while the hydrophobic stretch of the mouse gene occurs at the extreme amino-terminus, a more conventional location for cleaved signal sequences. Although portions of sequence N-terminal to the *Drosophila* signal sequences are conserved, suggesting a functional role, the mouse gene lacks this region.

The overall level of amino acid identity between *Hhg*-1 and *hh* carboxy-terminal to the signal sequences is 46%. A closer examination shows that the amino terminal portion, from residues 25 to 187, displays 69% identity, while remaining residues in the carboxy-terminal portion display a much lower 31% identity. Like *hh*, the *Hhg*-1 coding sequence

is divided into three exons, and the boundaries of these exons are at the same positions within coding sequence as those of the three *Drosophila hh* exons. Curiously, the boundary between coding sequences of the second and third exons occurs near the transition from high to low levels of overall sequence conservation. The coincidence of these two boundaries suggests a possible demarcation of functional domains within these proteins. This location within *Hhg*-1 coding sequence also coincides approximately with the site of a presumed proteolytic cleavage.

EXAMPLE 10
HUMAN CLONING OF *hh* GENES

Partial sequence for two human *hh* genes has been obtained by DNA sequencing of clones derived by PCR amplification from genomic DNA with *hh*-specific degenerate primers as outlined in Chang, *et al.*, *(Development*, 120:3339, 1994) and EXAMPLE 9 (FIGURE 11A and B). More extensive screening by the same approach, either with the same primers or with other primers from the *hh* coding region or with the human *hh* fragments seen in FIGURES 11A and B, is expected to yield at the least a third gene, and possibly more, since at least three genes are found in the mouse. These segments of human *hh* genes can be used to obtain full coding sequences for human proteins by the following cloning method commonly used by those of skill in the art and which are extensively described in the literature.

For example, ready-made cDNA libraries or RNAs from a variety of human sources, including various fetal stages and organs (from abortuses) and specific infant or adult organs (from pathological or autopsy specimens), are being tested for the presence of *hh* sequences by PCR or RT-PCR using the primers described in Chang, *et al.*, *supra*, and other primers derived directly from the sequence of the human fragments. Ready-made libraries containing *hh* sequences are being screened directly and, where necessary, new libraries are being constructed by standard methods from RNA sources containing *hh*

sequences. The probe for these screens is a mixture of all the distinct human *hh* fragments. Sequences of cDNA clones can then be determined. Most clones containing the probe sequences, which are located in the N region, will also include a full C coding region since standard methods of library construction result in cDNA clones that are most complete at their 3' ends. All full length *hh*-coding sequences obtained previously in vertebrates and invertebrates contain N and C sequences encoded in a single RNA. Screening is continued until complete open reading frames that correspond to all of the fragments of human *hh* genes are obtained. Specifically, 1.2×10^6 clones from a human fetal brain library (Stratagene, La Jolla, CA) was screened using a mixture of the two human *hh* fragments (FIGURE 11A and B) as probes. Twenty-nine clones were identified as specifically hybridizing with these probes.

Second, the RNA sources identified as containing *hh* sequences can be used as templates from anchored PCR (also referred to in the literature as RACE, for rapid amplification of cDNA ends). Briefly, this method provides a means to isolate further mRNA sequence in either the 5' or 3' direction provided that sequence is known from an internal starting point. Anchored PCR can also be used to isolate sequences from cDNA library.

Third, genomic libraries can be screened with the probes described in the first technique. Where necessary, human *hh* exons and coding sequences are being identified by hybridization to previously isolated human and mouse coding sequences by sequence determination, and by exon-trapping methods to identify all *hh* coding sequences within genomic clones; these coding sequences can be "stitched" together by standard recombinant DNA methods to generate complete *hh* open reading frames.

FIGURE 12 A and B show *in vitro* cleavage reactions of a *Drosophila hh* protein produced in *E. coli* and purified to homogeneity. This protein has residues 89-254 deleted, rendering it more soluble and easier to purify. It also contains a His_6 purification tag appended to the N-terminus. Autoproteolysis of this protein is triggered by the addition of reducing agents (DTT), and the resulting product corresponds to the C

fragment identified *in vivo*. FIGURE 12, Panel A shows a time course of cleavage after initiation by addition of DTT. Panel B shows incubations of concentrations ranging over three order of magnitude for a fixed time period (four hours), with no difference in the extent of conversion to the cleaved form. This concentration-independent rate of cleavage indicates an intramolecular mechanism of cleavage. Panel C shows the sequence around the cleavage site as determined by amino-terminal sequence of the cleaved fragment C. The cleavage site is denoted by the arrow, and the actual residues sequenced by Edman degradation of the C fragment are underlined. Panel C also shows an alignment of all published vertebrate *hh* sequences plus some of unpublished sequences from fish and Xenopus. The sequences shown correspond to the region of *Drosophila hh* where the cleavage occurs, and demonstrates the absolute conservation of the Gly-Cys-Phe sequence at the site of cleavage. Panel D shows a SDS-PAGE gel loaded with *in vitro* transcription/translation reactions as described in the previous Examples, using various *hh* genes as templates. dhh is *Drosophila*, *twhh* and zf*shh* are the *twiggy-winkle* and *sonic hh* genes of the zebrafish, and m*shh* is the *shh/Hgh-1/vhh-1* gene of the mouse. The translation mix included ^{35}S-labelled cysteine, used to visualize the resulting products by autoradiography. Note that each gene give a larger product (the precursor or U) and two smaller products of cleavage (N and C). The larger species is C for each of the vertebrate genes, whereas the *Drosophila* N is larger than C due to the presence of ~60 residues occurring amino-terminal to the signal sequence that are present in the vertebrate open reading frame. This panel shows that vertebrate *hh* proteins are processed similarly to the *Drosophila* protein. Panel E shows that Edman degradation of the C fragments releases ^{35}S counts on the first but not subsequent rounds for all these proteins, indicating that the site of autoproteolytic cleavage for all of these *hh* proteins is the amide bond to the amino-terminal side of the Cys residue that forms the center of the conserved Gly-Cys-Phe sequence highlighted in panel C. This is a generalizable approach to establish the composition of protein fragments from any other *hh* family members.

EXAMPLE 11

DIFFERENTIAL EXPRESSION OF TWO *hh* GENES IN AXIAL MESODERM AND IN NEURAL PROGENITORS.

Partial sequences corresponding to five distinct zebrafish *hh*-like genes were isolated and the complete coding sequences for two of these genes were obtained from an embryonic cDNA library. One of these two sequences is identical to that of the zebrafish *nhh-1* gene (Roelink, *et al.*, *Cell*, 76:761, 1994), and appears to correspond to the *shh* gene reported by Krauss, *et al.*, *(Cell*, 75:1431, 1993) (See FIGURE 13 description); the other gene, *tiggy-winkle* (Potter, B., The Tale of Mrs. Tiggy-Winkle, *The Penguin Group*, London, 1905), represents a novel vertebrate *hh*. Coding sequences for both are shown in alignment to mouse and chicken sequences of the *sonic/vhh-1* class (FIGURE 13b). Like other vertebrate *hh* homologues, the *twhh* and *shh* proteins contain an amino-terminal stretch of hydrophobic residues. These residues function as signal sequences since cleavage is observed when coding sequences are translated in the presence of microsomoses; vertebrate *hh* genes thus appear to encode secreted proteins, as previously reported for *Drosophila hh* (Kimmel C.B. & Warga, R.M., *Developmental Biology*, 124:269-280, 1987; Warge, R.M., & Kimmel, C.B., *Development*, 108:569-580, 1990).

The first four sequences were isolated from zebrafish genomic DNA (a gift from J. Pellegrino) using degenerate primers in polymerase chain reactions as described (Chang, *et al.*, *supra*). *twhh* and *shh* clones were isolated from a 20-28 hour cDNA library (a gift from R. Riggleman, K. Helde, D. Grunwald and J. Pellegrino) using the first three sequences as probes. The translational reading frames for *twhh* and *shh* were closed 12 and 16 codons, respectively, upstream of the putative initiating methionine.

Figure 13 shows the predicted amino acid sequences are shown in single letter code. 13(a) shows sequences common to five distinct *hh*-like genes are shown with a dot indicating identity with the corresponding residue of zebrafish *twiggy-winkle (twhh*; Potter 1905; *supra*), hh[zfB] and hh[zfC] is more diverged and appears to represent a

novel class. 13(b) shows amino acid sequences of *twhh* and *shh* are aligned to those of the *soniclvhh-1* class from chick and mouse (Riddle, *et al.*, *Cell*, 75:1401-1416, 1993; Chang, D.T., *et al.*, *Development*, *supra*; Echelard, Y., *et al.*, *Cell*, 75:1431-1444, 1993). Zebrafish *sonic hedgehog (shh)* is identical in sequence to *z-vhh-1* reported by Roelink, *et al.*, *Cell*, 76761-775, 1994. Based on expression and extensive sequence identity throughout most of the coding region, *vhh-1* and the *sonic* sequence reported here probably correspond to *shh* of Krauss, *et al.*, *Cell*, 75:1431-1444, 1993, diverges dramatically throughout a 26 residue stretch near the carboxy-terminus. Rat *vhh-1/sonic hh* (Roelink, *et al.*, supra.) was excluded in this alignment because of its 97% sequence identity to the predicted mouse protein. Residues identical in all four sequences are boxed, and a dash indicates a gap in the alignment. The arrow indicates the predicted signal sequence cleavage site (von Heijine, G. , *Nucleic Acids Res.*, 14, 4683-4690, 1986) for *twhh*. The amino-terminal hydrophobic stretch common to all four *hh* genes is shaded. 13(c) shows percent identity of residues carboxy-terminal to the hydrophobic region.

Figure 14 shows a comparative expression of *twhh*, *shh*, and *pax-2* during zebrafish embryogenesis. Whole mount *in situ* hybridizations on 0-36 hour embryos were performed using a modification of the procedure of Tautz and Pfeifle, *Chronosoma*, 98:81-85, 1989, with antisense probes. Transcript localization is revealed by the purple product of an alkaline phosphatase enzymatic reaction. Staging of the embryos is according to Westerfield, M., *(The Zebrafish Book*, University of Oregon Press, Eugene, 1993). Transcripts were visualized by *in situ* hybridization to whole embryos. (a, b) *twhh* expression in a single late shield stage embryo. (a) Dorsal view, animal pole is to the top. The triangular shape of expression is characteristic of axial mesoderm-forming cells of the hypoblast (Statchel, S.E., *et al.*, *Development*, 117:1261-1274, 1993). (b) Lateral view: the thicker layer of cells on the left (dorsal) side of the embryo is the embryonic shield; the two arrows indicate the *twhh*-expressing hypoblast cells and the non-expressing epiblast. Antcrior is to the left in all subsequent embryos. Dorsal is to the top in all lateral views. (c, d) A single embryo at the end of gastrulation (100% epiboly)

with *twhh*-expressing cells. (d) Caudal-dorsal view. Note the wide patch of stain in the presumptive tailbud which narrows anteriorly. (e, j) Early somitogenesis (11.5 hour, 3-4 somite) embryos; optic vesicles have not begun to evaginate from the wall of the diencephalon. (e, h, k) Lateral views of developing brain. (f, i, l) Dorsal views of developing brain. (e, f, g) Localization of *twhh*-expressing in a single row of cells that will form the flood plate. The arrowhead marks a parch of *twhh*-expressing cells lateral to the tailbud. (h, i, j) Localization of *shh*. *shh* is also expressed strongly in the protuberance. (j) Lateral view of the developing tail. *shh* is also expressed strongly in the protuberance. (j) Lateral view of developing tail. *shh* is expressed in cells that will form both floor plate and notochord. (k, l, m) Localization of *pax-2* during early optic vesicle formation; (m) also shows *twhh* expression. (k) 12 hour (4-5 somites) embryo. (l) 12.5 hour (5-6 somites) embryo. Expression of *pax-2* in the developing optic vesicle is in a gradient away from the protuberance. Note the expression of *pax-2* (asterisk) at the future midbrain-hindbrain border. (m) *twhh* (arrow) and *pax-2* expression in a 6-7 somite (13 hour) stage embryo. Note differential expression of *twhh* in ventral neural keel (corresponding to neural tube in other vertebrates). (n-s) Embryos at end of somitogenesis (22-24 hours). (n, o, p) Localization of *twhh*. (n, o) Developing brain. Note isolated groups of cells staining in the diencephalon (filled triangles) and the protuberance (arrowhead), and floor plate expression underlying the midbrain and hindbrain. The floor plate expression is contiguous caudally along the axis. (n) Lateral view. (o) Dorsal view. (p) Lateral view of tail. Expression is restricted to the floor plate. (q, r, s) Localization of *shh*. (q, r) Developing brain (q) Lateral view. *pax-2* expression in the otic vesicle is indicated. (r) Dorsal view. Expression in the protuberance (arrowhead) and in the neural keel. (s) Lateral view of tail. Expression is strongest in the floor plate, but contrary to the report of Krauss, *et al.*, supra., is still also in the notochord. Abbreviations: white e - epiblast; h - hypoblast; tb - tailbud; p - protuberance; c - eye; ov - optic vesicle; ot - otic vesicle; fp - floor plate; nc - notochord; asterisk - midbrain-hindbrain boundary or *pax-2*-labeled prospective midbrain-hindbrain boundary; t - telencephalon.

Comparison of *twhh* and *shh* expression patterns (Krauss, *et al, supra*), reveals that both gene are predominantly expressed in midline structures, albeit with notable differences in regard to timing, rostra-caudal extent, and tissue restriction. Expression of *twhh* is first detected during gastrulation in the dorsal mesoderm (FIGURE 14a, b); this expression occurs in a band corresponding to a subset of the embryonic shield, a structure, analogous to Spemann's organizer in *Xenopus* (Stachel, *et al., Dev.*, 117:1261-1274, 1993, and reference therein; Ho., R., *Seminars in Developmental Biology*, pg.3, 1992). In concert with the movements of convergence and extension, this band of *twhh* expression shortens along the equatorial plane and extends along the incipient embryonic axis until, by the end of gastrulation, expression occurs throughout the entire axis (FIGURE 14c,d). Early in somitogenesis, *twhh* RNA is found restricted to presumptive ventral neural tissue along the entire body (FIGURE 14e, f, g), the only exception being cells in and near the tailbud (FIGURE 14g). In contrast to the neural restriction of *twhh*, *shh* is localized both to presumptive neural and notochordal cells (FIGURE 14j).

As somitogenesis proceeds, ventral midline expression of *shh* and *twhh* is reduced in most of the prospective forebrain, but remains strong in an anterior patch of midline cells within the floor of the prospective diencephalon (FIGURES 14e, f, for *twhh*; FIGURES h, i for *shh*), This patch later will give rise to the protuberance (Schmitt, E.A. and Dowling, J.D., *J. Comp. Neur.*, 344:532-542, 1994), an anterior extension of the diencephalon. This structure, which is medial and just rostral to the developing optic stalks, is the site we propose as the focus of early patterning activity for the developing eyes (see below). By the end of somitogenesis, both *twhh* and *shh* are strongly expressed in the floor plate (FIGURES 14p, s), although *shh* transcripts remain detectable in the notochord at this stage and at 36 hours of development (FIGURES 14s; later stage not shown). At 28 hours, *twhh* transcripts are also found in a small cluster of cells within the first gill arch (not shown), as also reported for *shh* at 33 hours of development (Krauss, *et al., supra*).

Differences between *twhh* and *shh* expression are apparent from the beginning of gastrulation, since *twhh* RNA can be detected as early as the shield stage while *shh* is first detected later, at about 60% epiboly (not shown; (Krauss, *et al.*, *supra*). In addition, *twhh* transcripts are restricted to neural tissues early in development, and are never detected in the notochord (compare FIGURE 14g to FIGURE 14j). Later differences in expression include differential rostra-caudal restriction within the diencephalon and midbrain and weaker and more restricted expression of *twhh* in the protuberance (compare FIGURES 14n and 14q), such that the later domain of *twhh* expression in the brain appears to constitute a subset of the *shh* domain. In addition, *shh* but not *twhh* is expressed in the developing fin bud (Krauss, *et al.*, *supra*). Comparison of *shh* and *twhh* expression patterns to this previously reported for *hh* homologues in zebrafish and other vertebrate species indicates that *shh* is the zebrafish homologue of the *sonic/vhh-1* class while *twhh* represents a novel class of vertebrate *hh*.

EXAMPLE 12

DEVELOPMENTAL CONSEQUENCES OF ECTOPIC hh EXPRESSION DURING ZEBRAFISH EMBRYOGENESIS

To gain insight into the potential roles of *hh* products in development, synthetic *twhh* and *shh* mRNA was injected into 1-8 cell embryos. This technique yields a mosaic but fairly uniform pattern of expression, as determined for the control mRNA encoding β-galactosidase (not shown). Uniformity of expression is in good agreement with fate mapping studies of the early zebrafish embryo (Kimmel & Warga, *supra*; Warga & Kimmel, *supra*; Heide, *et al.*, *Science*, 265:517-520, 1994), which indicate that blastomeres undergo extensive cell mixing during the cleavages prior to gastrulation. We note that mosaicism of expression caused surprisingly little variation in the phenotypes of the *hh* injected embryos, possibly due to secretion of *hh* gene products.

Embryos injected with synthetic *twhh* or *shh* mRNA (*hh* RNA) exhibited numerous yet highly reproducible abnormalities in comparison to control embryos injected with *lacZ*

mRNA. These abnormalities, discussed below, are primarily defects in the brain and eyes. Although the effects of ectopic *twhh* and *shh* expression were qualitatively similar, the incidence and severity were greater with *twhh* RNA (see text below, FIGURE 15 and FIGURE 16). The proteins encoded by these two genes have qualitatively similar biological activities, but apparent differences in potency.

FIGURE 15 shows the effects of ectopic *hh* on zebrafish development. Wild type zebrafish, *Danio rerio*, Ekkwill Waterlife Resources) were maintained at 28.5°C, some embryos were then cultured overnight at RT. Zebrafish embryos were injected at the 1-8 cell stage with *twhh*, *shh*, or *lacZ*RNA and examined at 28 h of development. (a-c) Dorsal view of the midbrain-hindbrain region; anterior is left. (a) *lacZ*. (b) *twhh*. (c) *shh*. (d-f) Frontal optical section of the forebrain region; anterior is up. (d) *lacZ*. (h) *twhh*. (f) *shh*. (g-1) Lateral view of the eye region; anterior is left. (g) *lacZ*. (h) *twhh*. (i) *twhh*. At levels caudal to the prospective brain, the notochord, somites, and neural keel formed by most *hh*-injected embryos appeared grossly normal except for an overall shortening and dorsal curvature of the axis. A minority of *hh*-injected embryos (15% are not shown) displayed partially bifurcated axes, containing duplicated axial mesoderm and parallel neural keels, each neural keel comprising ventral midline cells and some bilaterally symmetric lateral cells (not shown). Although we have not determined the primary cause of these axial defects, analysis of late gastrulation stage embryos suggests hat the bifurcation may result from difficulties in epiboly and convergence. Abbreviations: mv - mesencaphalic ventricle; rv - rhombencephalic rentricle; asterisk - midbrain-hindbrain boundary; ot - otic vesicle; tv - third (diencephalic) ventricle; r - retina or retina-like structure; l - lens or lens-like structure; pe - pigmented retinal epithelium.

Morphological defects in the brain and other rostral neural derivatives occur at high frequency in *hh*-injected embryos. The three ventricles of the fish brain normally apparent at 28 hours of development - the rhombencephalic, mesencephalic (FIGURE 15a), and diencephalic (third ventricle; FIGURE 15d) - are not formed in the brains of *hh* injectees (FIGURES 15b, c; FIGURES 15e, f)), despite the obvious presence of a lumen.

The prominent construction normally present at the midbrain-hindbrain boundary also is absent (compare FIGURE 15a to FIGURES 15b, c). Formation of this constriction requires function of *pax-2* (Krauss, *et al.*, *Nature*, 353:267-270, 1991; Krauss, *et al.*, *Nature*, 360:87-89, 1992), which normally is expressed in a band at the midbrain-hindbrain boundary (Krauss, *et al.*, *supra*; Krauss, *et al.*, *Development*, 113:1193-1206, 1991) *pax2* expression at this boundary is not disrupted by *hh* RNA injection, however, indicating that this phenotype does not result from disruption of rostra-caudal information.

Defects in eye development also occur at high frequency in embryos injected with *hh* RNA. Thus, while at 28 hours the normal zebrafish eye has a lens and a retina with pigmented epithelium (FIGURE 15d, g), *hh*-injected embryos usually fail to develop lenses and retinal pigmentation (FIGURE 15e, h). Eye duplications are also observed at low frequencies (FIGURE 15i). The poorly developed eyes do not appear to result from a simple delay in development since pigmentation elsewhere in injected embryos appears in its normal time course. Examined at three days of development, the consequences of *hh* RNA injection include defects that range from complete absence of eyes to partially formed eyes lacing a ventral portion of the retina.

The eye phenotypes caused by *hh* RNA injection resemble those produced by treatment of zebrafish and *Xenopus laevis* embryos with retinoic acid. In *Xenopus*, phenotypes range from reduction of the eye and absence of the lends to eyes with retinal folds (resembling duplicated dyes) and multiple small lenses (Manns, M. & Fritzsch, B., *Neurosci. Lett.*, 127:150-154, 1991). In zebrafish, exposure to retinoic acid during gastrulation interferes with the formation of the eye (Holder, N. & Hill, J., *Development*, 113:1159-1170, 1991), while exposure during formation of the optic primordia induces formation of duplicated retinas and extra lenses (Hyatt, *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:8293-8297, 1992). Patterning effects of retinoic acid upon the developing chick limb appear to be mediated through ectopic activation of the endogenous *sonic hh* gene (Riddle, *et al.*, *supra*), these results with ectopic *hh* expression suggest the possibility of

a similar mechanism underlying the patterning effects of retinoid acid treatment in the vertebrate eye.

EXAMPLE 13

hh EXPRESSION IN THE OPTIC VESICLE SPECIFIES PROXIMAL FATES AT THE EXPENSE OF DISTAL FATES

To further elucidate the role of *hh* in eye development we utilized *pax-2* and *pax-6* (Krauss, *et al.*, *EMBO J.*, 10:3609-3619, 1991; Pitischel, *et al.*, *Development*, 114:643-651, 1992) were utilized as positional markers to examine the effects of ectopic *hh* expression on the optic vesicle. As the optic vesicle evaginates from the lateral walls of the zebrafish forebrain (Schmitt, E.A. & Dowling, J.D., *J. Comp. Neur.*, 344:532-542, 1994), *pax-2* is expressed in a gradient, with highest RNA levels in the anterior and ventral regions of the optic vesicle (Krauss, *et al.*, *supra*; FIGURE 14k, l, m). Immediately adjacent to the maximum of this *pax-2* expression gradient is the region of the dicnecphalon termed the protuberance (Schmitt & Dowling, *supra*), where both *twhh* and *shh* but not *pax-2* are strongly expressed (FIGURES 14e, f, h, i, m). The concentration gradient of *pax-2* expression in the eptic vesicle thus appears to incline downward from its maximum at a location adjacent to the site of *twhh* and *shh* expression in the protuberance. Superposition of developmental fate within the optic vesicle (Schmitt, *et al.*, *supra*), upon the pattern of *pax-2* expression suggests that the gradient of *pax-2* RNA prefigures the future proximal/distal axis of the eye.

Ectopic *hh* alters the expression of *pax-2*, *pax-6*, and *F-spondin*. Zebrafish embryos were injected at the 1-8 cell stage with *twhh* or *shh* RNA and the pattern of *pax-2*, *pax-6*, or *F-spondin* expression was examined by whole mount *in situ* hybridization. Control embryos injected with *lacZ* RNA were performed in every case and displayed wild-type expression patterns. At embryo stage, the anterior-posterior axis of the optic vesicle corresponds to the future proximal-distal axis of the eye. During the next hour of

development, the posterior edge of the optic vesicle will separate from the diencephalon (Schmitt and Dowling, *Comp. Neur.*, 344:532-542, 1994).

Injection of either *hh* RNA causes uniform initiation of *pax-2* expression along both the proximal-distal and dorsal-ventral axes of the optic vesicle as it begins to evaginate. The ectopic *pax-2* expression appears at the same time as normal *pax-2* expression is initiated in the eye, and in some cases, is also seen in the diencephalon between the optic vesicles. At the end of somitogenesis, a time when *pax-2* would normally be restricted to the optic stalk, *pax-2* RNA in *hh* injected embryos is detected in all but the most distal portion of the optic vesicle.

The effects of ectopic *hh* on expression of *pax-6*, which encodes a transcription factor critical for eye development was also studied. At 22 hours of zebrafish development, *pax-6* is normally expressed in the lens and in most of the distal part of the optic cup (Krauss, *et al.*, *supra*; Puschel, *et al.*, *Development*, 114:643-651, 1992). In *hh*-injected embryos, *pax-6* is repressed in the optic vesicle, although many embryos retain *pax-6* expression in the most distal cells. With regard to *pax-2* and *pax-6* as markers of positional identity, *hh* expression in the optic vesicle can be characterized as inducing proximal fates and repressing distal fates.

The distal part of the optic vesicle is the most refractory to *hh*-induced changes in both *pax-2* and *pax-6* gene expression. Due to a later rotation, this distal portion of the optic vesicle will give rise to the dorsal portion of the mature eye (Schmitt, *et al.*, *supra*); interestingly, this is the portion of the eye that remains in 3-day old injected embryos with intermediate phenotypes (see above).

Lesions in the *pax-6* gene have been assigned as the basis for the *Aniridia* (Ton, *et al.*, *Cell*, 67:1059-1074, 1991; Glaser, *et al.*, *Nat. Genetics*, 2:232-239, 1992), *Small eye* (Hill, *et al.*, *Nature*, 354:522-525, 1992), and *eyeless* mutations (Quiring, *et al.*, *Science* 265:785-789, 1994), in humans, mice and *Drosophila*, respectively; *pax-6* function thus

appears to be critically required for eye development in *Drosophila* and mammals. As we argue here, *hh*-encoded activities also appear to play a role in vertebrate eye development, and this suggests a further molecular parallel between vertebrates and insects, since the role of *hh* in *Drosophila* eye development is well established (Mohler, *et al., supra*; Ma, *et al., supra*; Heberlein, *et al., supra*; Lee, *et al., supra*). The reciprocal and non-overlapping patterns of *hh* and *pax-6* expression in the developing *Drosophila* eye (Ma, *et al., supra*; Quiring, *et al., Science*, 265:785-789, 1994), suggest the possibility of *pax-6* repression by *hh*, but whether *hh* functions by similar mechanisms in vertebrate and *Drosophila* eye development is a questions that requires further investigation.

In mice, the dosage of *pax-6* protein is crucial for normal eye development (Hill, *et al., supra*). *Small* eye heterozygotes develop an abnormally small lens (Hogan, *et al., J. Embryol. Exp. Morph.*, 97:95-110, 1986; Hogan, *et al., Development*, 103 Suppl., 115-119, 1988), as do *hh*-injected embryos with weaker phenotypes (FIGURE 14f). *Small* eye homozygotes lacking lenses eventually generate and the animals lack eyes at birth (Hogan, *et al., supra*; Hogan, *et al., supra*), as do many of the *hh*-injected embryos at three days of development. These parallels suggest that many of the later eye defects observed in *hh*-injected zebrafish may be caused by partial or complete repression of *pax-6* during eye development.

EXAMPLE 14

GENETIC ABLATION OF *hh* FOREBRAIN EXPRESSION CAUSES LOSS OF PROXIMAL FATES IN THE OPTIC VESICLE

The patterns of *twhh* and *shh* expression (FIGURE 14) and the effects of ectopic *hh* expression (FIGURE 15) are consistent with a normal role for *shh* and *twhh* in eye development. If *hh* activities indeed play a normal role in promoting proximal fates within the developing eye, removal of *hh* activities would be expected to result in a loss of proximal fates. In embryos homozygous for the *cyclops* mutation ventral neural

structures fail to form and the developing eyes fuse at the midline,, yielding an embryo with a single eye (Hatta, *et al.*, *Nature*, 350:339-341, 1991). The missing ventral structures in *cyclops* mutants include the regions where we observe expression of *twhh* and *shh*, and we therefore examined the effects of the *cyclops* mutation on *hh* expression.

cyc^{b16} (Hatta, *et al.*, *Nature*, 350:339-341, 1991), heterozygous adults (a kind gift of R. Riggleman) were spawned and their offspring analyzed by whole mount *in situ* hybridization. Detection of *pax-2* and either *twhh* or *shh* RNAs in embryos homozygous for the *cyc* mutation or their wild-type siblings. *twhh* RNA is only expressed in the presumptive tailbud (caret) of *cyc* embryos. As reported by Krauss, *et al.*, *Cell, supra*, neural expression of *shh* is abolished in *cyc* embryos. Strong *pax-2* expression was observed in the optic vesicles of wild-type embryos which is significantly reduced in *cyc* mutant embryos.

twhh RNA in *cyclops* embryos is found only in a small patch of cells at the presumptive tailbud and neural expression was not detected at any later stage examined. Neural expression of *shh* is also lost in *eye* mutants, although expression in the notochord is reunited (Krauss, *et al.*, *supra*; data not shown).

Since the *eye* mutation appears to ablatc *hh*-expressing cells in the developing brain, this mutation can be used as a genetic tool to examine the requirement for *hh* function in eye development. Iiatta, *et al.*; Hatta, *et al.*, *Proc. Natl. Acad. Sci. USA*, 91:2061-2065, 1994), recently demonstrated that *pax-6* expression is fused at the midline due to loss of ventral midline cells that normally do not express *pax-6* and, in addition, *pax-2* expression in the fused eye of *eye* mutant embryos is reduced. We extended these observations to an earlier stage when the optic vesicles first form and found that *pax-2* expression is weak and fails to extend within the vesicles in *eye* mutants. In conjunction with the results of ectopic *hh* expression, these observations suggest that *hh* signaling that activity promotes and is required for the induction of proxima fates within the eye vesicle.

In this model, we propose that the protuberance acts as a proximal patterning center for the developing zebrafish eye by providing a localized source of *hh* activity.

EXAMPLE 15

hh ACTIVITY VENTRALIZES THE DEVELOPING BRAIN

Previous work has established an important role of signals from the floor plate and notochord in ventral patterning of the neural tube (Jessell, T.M., & Dodd, J., *Cell*, 69:95-110, 1992). For example, Goulding, *et al.*, *Development*, 117:1001-1016, 1993, recently demonstrated that notochord and floor plate grafts can repress the normal lateral expression of *pax-6* in the neural tube. Other recent work has implicated *hh* activity in at least some aspects of ventral neural tube patterning (Echelard, *et al.*, *Cell*, 75:1417-1430, 1993; Krauss, *et al.*, *supra*; Roelink, *et al.*, *supra*); consequently, we examined *hh*-injected embryos for effects on *pax-6* expression in the brain.

In the zebrafish at 22 hours of development, *pax-6* is expressed in dorso-lateral regions of the diencephalon and in a ventro-lateral domain of the hindbrain and spinal cord that excludes the floor plate and adjacent cells (Krauss, *et al.*, *supra;* Puschel, *et al.*, *supra*). This pattern of expression is reciprocal to that of both *twhh* and *shh* in the diencephalon (compare FIGURES 14q and 14i) and in the hindbrain. *hh* RNA injection caused repression of *pax-6* in the more ventral domain in the diencephalon, while more dorsal expression persisted. In addition, *pax-6* expression was significantly reduced ventrally in rhombomeres 1, 2, and 4 and, in some cases, was completely abolished in these rhombomeres. The repressing effect of ectopically expressed *hh* and *pax-6* in normal embryos are due to repression of *pax-6* by nearby *hh* expressing cells.

Since absence of *pax-6* expression is a feature of the ventral midline, repression of *pax-6* in lateral positions suggests ventralization. Consequently, *twhh* was injected into embryos for analysis of induction of a floor plate marker, *F-spondin* (Riddle, *et al.*, *supra*). As described above, ectopic *twhh* induces *F-spondin* expression at more dorsal

levels in the midbrain and anterior hindbrain. The effects of *hh* upon expression of both *pax-6* and *F-spondin* indicate a ventralization of the brain. Adoption of ventral cell identity by lateral cells might explain their failure to form ventricles (FIGURE 15a-f).

The ventralizing activities of *twhh* confirm and extend those previously reported for *shh/vhh-1* class genes of chicken, zebrafish, and rat (Echelard, *et al.*, *supra*; Krauss, *et al.*, *supra*; Roelink *et al.*, *supra*). The early restriction of *twhh* to midline neural progenitors, however, suggests that it may play a specific role in the homeogentic mechanisms of floor plate maintenance and expansion (Placzek, *et al.*, *Dev.*, 117:205-218, 1993). In the zebrafish, wild type cells in *cyclops* hosts can contribute to and induce adjacent cells to form floor plate, but only when the transplanted cells populate the neural plate and not the notochord (Hatta, *et al.*, *Nature*, 350:339-341, 1991). We have demonstrated that, in *cyclops* mutants, midline expression of *twhh* is lost while *shh* expression is maintained in the notochord (FIGURE 18; Krauss, *et al.*, *supra* for *shh*); taken together, these results suggest that the homogenetic floor plate signal lost in the *cyclops* mutant may be encoded by the *twhh* gene. In the chick and rat, the floor plate retains auto-inductive potential long after the loss of floor plate inducing properties by the notochord, despite continued expression of *shh/vhh1* in the notochord (Roelink, *et al.*, *supra*; Placzek, *et al.*, *supra*; Yamada, *et al.*, *Cell*, 73:673-686, 1993). Although no homologues of the *twhh* class have been reported in other vertebrates, expression of other *hh* homologues in patterns more like those of *twhh* might help explain these discrepancies.

EXAMPLE 16

TWO DISTINCT SIGNALING PROTEINS DERIVE FROM THE *twhh*-ENCODED PRECURSOR

Endogenous *hh* protein in *Drosophila* is fund predominantly as an amino- and a carboxy-terminal fragment (N and C, respectively) derived by an internal auto-proteolytic

cleavage of a larger precursor (U for uncleaved), which also occurs *in vivo* but at lower levels (Lee, *et al.*, *supra*). Determinants within the amino-terminal domain appear not to be required for auto-proteolytic activity, whereas mutations affecting the carboxy-terminal domain can block auto-proteolysis and reduce activity *in vivo* (Lee, *et al.*, *supra*). The auto-proteolysis is blocked by a substitution of alanine for the histidine normally present at position 329. This histidine is absolutely invariant in alignments of all known *hh* genes, and its sequence context suggests a catalytic role in auto-proteolysis (Lee, *et al.*, *supra*).

FIGURE 17 shows zebrafish *twiggy-winkle hedgehog* derivatives. 17(a) Cartoons of various *twhh* open reading frames. SS (shaded) is the predicted N-terminal signal sequence for secretion of these proteins and encompasses the first 27 amino acids of each open reading frame. The arrow indicates the predicted internal site of auto-proteolytic cleavage. Amino acid residue numbers are according to Figure 13b. The filled triangle denotes the normal termination codon for the *twhh* open reading frame. Construct U_{HA} contains a mutation that blocks auto-proteolysis (the histidine at residue 273 is changed to an alanine; see Lee, J.J., *et al.*, *supra*.). Construct $U356_{HA}$ contains a stop codon in place of amino acid residue 357 as well as the H273A mutation in U_{HA}. Construct N encodes just the first 200 amino acids of *twhh*. Construct C has had the codons for residues 31-197 deleted. 17(b) shows *in vitro* translation of the expression constructs shown schematically in part a. Constructs were translated *in vitro* in the presence of ^{35}S methionine and analyzed by autoradiography after SDS-PAGE. The protein products are shown schematically to the left. Lanes 1 and 6: Auto-proteolysis of the full-length (U_{SS}) protein creates two fragments, an N-terminal fragment (N_{SS}) and a C-terminal fragment (C). Lane 2: Construct U_{HA} only makes an uncleaved form of *twhh* protein that comigrates with U_{SS} *twhh* via auto-cleavage. Lane 5: Construct C encodes processed and unprocessed forms which are visible as two bands migrating closely together. The bottom band is the C protein made from auto-proteolysis of the U_{SS} (Δ31-197). All constructs were made by *in vitro* mutagenesis of expression construct T7TS*twhh* (see FIGURE 15) using the method of RPCR. The sequence of all constructs were confirmed

by dideoxy sequencing. *In vitro* translations were performed according to manufacturer's instructions (Promega).

The vertebrate *hh* proteins encoded by *shh, twhh* and mouse-*shh/Hhg-1* also undergo auto-proteolysis to yield two smaller species from a single larger precursor (Lee, *et al., supra*; Chang, *et al., supra*; see lanes 1 and 6 in FIGURE 17b). The invariant histidine to alanine mutation to generate a construct encoding a form of the *twhh* protein that is not auto-proteolytically cleaved (U_{HA}). We have also introduced a nonsense codon and deleted a segment of coding sequence to generate constructs that produce either the amino- or the carboxy-terminal domains of *twhh* (N and C, respectively; see lanes 4 and 5 in FIGURE 17b); constructs are schematically diagrammed in FIGURE 17a). To target these proteins to the secretory pathway, all constructs retained the normal *twhh* signal sequence.

Synthetic mRNAs transcribed from these constructs were injected to examine the role of processing and to assay the activities of individual protein fragments; the results are summarized in Table I and are based on the activities presented in FIGURE 15. The most striking conclusion from these experiments is that N and C both exhibit activity, and that these activities are distinguishable. Thus, although both N and C are capable of ectobpically activating *pax-2* in the developing eye, thereby providing an internal injection control, only N was capable of efficiently repressing *pax-6* (FIGURE 16). Later effects on lens development were also more extreme for N, consistent with the role of *pax-6* in lens development suggested by its mutant phenotypes in mice. (See Ton, C.C., *et al., Cell* 67:1059-1074, 1991; Glaser, T., *et al., Nat. Genetics* 2:232-239, 1992; Hill, R.E., *et al., Nature* 354:522-525, 1991; Hogan, B.L., *et al., J. Embryol. Exp. Morph.*, 97:95-110, 1986; and Hogan, B.L., *et al., Development*, 103Suppl.:115-119, 1988.)

In considering the activity of delta N-C, it is important to recognize the activity of endogenous *hh* genes in these experiments, which are inhibited by delta N-C and fragments thereof. (see Example 18 and FIGURE 18 for further discussion)

The uncleaved U_{HA} protein is only somewhat less active than C in inducing *pax-2*, but it also was not able to repress *pax6* efficiently (FIGURE 16). The latter is particularly notable since the U_{HA} protein (U35$_{HA}$; see FIGURE 17a, b) has activities not significantly different from N (FIGURE 16). Thus, in addition to carrying determinants important for auto-proteolysis and *pax-2*induction, the C-terminus also contains a domain inhibitory to N-terminal function when in the context of the uncleaved *hh* protein. The C-terminus can also inhibit N action by an intermolecular mechanism (Lai, *et al.*, *supra*). The existence of such an inhibitory domain in C suggests that if autoproteolyis can be modulated, such modulation might regulate the activity of *hh in vivo*. This possibility highlights the importance of ascertaining the processed state of *hh* proteins expressed in any particular patterning center to understand the potential *hh* activities generated.

EXAMPLE 17

DUAL ROLES OF *hh* SIGNALING PROTEINS IN EARLY EYE AND BRAIN PATTERNING

In understanding the normal roles of N and C in eye and brain patterning, the N and C derivatives of the *Drosophila hh* gene may offer some insight. The *Drosophila* N derivative is retained close to its embryonic site of synthesis in a segmentally striped pattern (Tabata and Kornberg, *Cell*, 76:89-102, 1994; Taylor, *et al.*, *Mech. Dev.*, 42 89-96, 1993), is cell-associated when expressed in cultured cells, and is effectively bound by heparin agarose *in vitro*, suggesting the possibility of extracellular matrix association. The C-terminal fragment, in contrast, is not bound effectively by heparin agarose, is almost quantitatively released into the culture supernatant of expressing cultured cells, and is only diffusely localized in embryos. Although the activities of individual fragments have not been assayed, the biochemical differences and tissue distributions of *Drosophila* N and C may account for the short and long range nature of the functions associated with *hh* during *Drosophila* development.

Although the tissue distributions of zebrafish N and C are not known, their activities in ectopic expression assays are also suggestive of short- and long-range functions when considered in the context of normal expression patterns of *hh*, *pax-2* and *pax-6*. The normal gradient of *pax-2* expression in the optic vesicle extends a substantial distance from its maximum adjacent to the site of *hh* expression in the protuberance; the ability of ectopic C to activate *pax-2* therefore suggests that, consistent with the distribution of C in *Drosophila*, zebrafish C may carry out a long-range function. Repression of endogenous *pax-6* expression, in contrast, appears to be a short-range function since *pax-6* expression occurs close to endogenous *hh* expression. Efficient repression of *pax-6* is an attribute of constructs producing N, and a short-range function for N would be consistent with the distribution of N in *Drosophila*.

Two types of *hh*-dependent activity have been reported for *hh*-transfected cultured cells. One is the apparent contact-dependent induction of floor plate markers (Roelink, H., *et al.*, *Cell* 76:761-775, 1994); the second induction of sclerotome markers in presomitic mesoderm, is diffusible and acts at long-range.

EXAMPLE 18

CHARACTERIZATION OF XENOPUS *hh*

1. Materials and Methods

cDNAs encoding full-length Xenopus hedgehogs, or encoding amino terminal or carboxy terminal domains linked to secretory leader sequences were transcribed *in vitro* to yield translatable messenger RNA. The synthetic messenger RNAs, and control mRNAs, were microinjected into the animal poles of cleavage stage Xenopus embryos, which were allowed to develop to the blastula stage, at which time the animal cap explants were prepared from the upper one fourth of the embryo. These blastula cap explants were then cultured *in vitro* in physiological saline in the presence or absence of the transforming growth factor beta family member, recombinant human activin A. All explants were allowed to develop until control embryos had grown to neurula stage, or to tadpole stage. Importantly, blastula caps left untreated differentiate from ectoderm into atypical epidermis. Blastula caps treated with activin differentiate into mesodermal and neural cell types. Thus, the question was whether hedgehog, or its proteolytic derivatives, would change the differentiation of cells away from becoming epidermis, and into another cell type. A second question was whether hedgehog can work with activin to alter the normal response of the tissue to either factor by itself.

Explants were then extracted to yield mRNA by methods commonly used by those of skill in the art, which was used as template with reverse transcriptase to yield cDNA. The cDNA was then used as template with various sets of primers for PCR for specific genes, reverse-transcriptase-polymerase chain reaction, or RT-PCR. This results in specific

amplification of radioactive products which are diagnostic for the presence and level of the messenger RNAs which were present in the explants. Samples were separated on polyacrylamide gels, which were exposed to X-ray film to yield the bands shown in the figures. Thus, the darker bands correspond to a greater level of the specific mRNA.

FIGURE 18A and B demonstrate that hedgehog induces pituitary and anterior brain genes, and can cooperate with activin or with neural inducers such as noggin and follistatin which are induced by activin to elevate expression of these genes in explanted embryonic tissue. All odd numbered lanes lack reverse transcriptase in the RT-PCR reaction and are negative controls. All even numbered lanes have this enzyme, and thus give specific bands to mRNA. In Panel A, Lanes 1-2 are control blastula caps, lanes 3-4 are Xenopus hedgehog-expressing blastula caps, lanes 5-6 are control blastula caps treated with activin, lanes 7-8 are hedgehog-expressing blastula caps treated with activin, and 9-10 are prolactin-expressing blastula caps treated with activin to serve as a control for simply expressing a secreted protein in the blastula cap. The primers used for the assay are shown to the left of each panel, *i.e.*, XAG 1 is a cement gland marker, XANF1B is a pituitary marker, otx-A is an anterior brain marker, en-2 is a midbrain-hindbrain boundary marker, krox 20 is a rhombomere-specific hindbrain marker, HIHbox 6 is a posterior hindbrain marker, NCAM is a general neural marker, activin is a control for mesoderm induction, and elongation factor is a positive control to shown that all even numbered lanes did in fact have cDNA present.

The panel labelled XANF1B detects a pituitary gene. Lane 4 (panel A) shows that hedgehog induces this pituitary marker, and thus likely pituitary cell types, in blastula cap explants (see also FIGURE 20, lane 6, for a stronger signal showing this), when compared to control explants in the absence of hedgehog (lane 2), which do not express this gene. Lane 6 shows that explants treated with activin, in the absence of hedgehog, also express the pituitary gene. Lane 8 shows that explants treated with both hedgehog, and with activin, give highest levels of the pituitary gene. Lane 10 proves that this effect

of hedgehog is specific, since prolactin, another secreted protein, does not lead to this elevated level of pituitary gene.

The panel labelled OTX-A detects this anterior brain gene. Lane 4 (and 6 in Figure 20) shows that hedgehog can induce this neural-specific gene. Lane 8 shows that the level of this neural gene is highest in tissue treated with both activin and hedgehog, relative to hedgehog alone (lane 4), or activin along (lane 6), and control explants do not express this gene (lane 2). Again, this effect is specific to hedgehog, since prolactin (lane 10) did not lead to elevated expression of this gene. The panel labelled XAG-1 detects a cement gland-specific gene, and lane 4 shows that hedgehog induces this gene at high level.

In panel 18B, embryos were injected with N or ΔN-C, and some animal cap explants were treated with activin before culturing until sibling embryos reached tailbud stage. Lanes 1, 2: control animal caps from uninjected embryos. Lanes 3, 4: control animal caps from uninjected embryos, treated with activin. Lanes 5, 6: animal caps from embryos injected with N and treated with activin. Lanes 7, 8: animal caps from embryos injected with ΔN-C and treated with activin. Whereas N displays activities in activin-treated explants similar to those of X-*bhh* (see B) ΔN-C produces the opposite effect, decreasing anterior and increasing posterior neural marker expression. As shown in Figure 18B, N behaves like X-*bhh* in that it induces elevated levels of XANF-2 and Otx-A (lane 6) relative to control activin-treated animal caps (lane 4). Moreover, N also leads to a decrease in the expression of more posterior markers, such as *krox*-20 and X*lHbox*-6, as observed following injection of X-*bhh*. In contrast to the activity of N (Fig. 4C, lane 6), ΔN-C *decreases* the expression of the anterior neural genes XANF-2 or Otx-A (Fig. 4C, lane 8) in activin-treated animal caps when compared to uninjected controls (lane 4). Moreover, ΔN-C also leads to an *increase* in the expression of more *posterior* markers, such as *En*-2 and X*lhbox*-6.

FIGURE 19 shows X-*bhh* modifies the anteroposterior pattern of neural gene expression in explants under the influence of endogenous neural inducers. (A) Isolation of dorsal

explants from injected embryos for the preparation of Keller sandwiches (Keller and Danilchik, 1988; Doniach, *et al.*, 1992; redrawn from Doniach, 1993). (B) Keller sandwiches were made from uninjected (lanes 1 and 2) and X-*bhh*-injected (lanes 3 and 4) embroys, total RNA was isolated when control embryos reached stage 20, and RT-PCR was used to analyze the expression of XAG-1 and neural markers. XAG-1 is a cement gland marker, XANF-2 is an anterior pituitary marker, Otx-A is a forebrain marker, *En*-2 demarcates the midbrain-hindbrain boundary, Krox-20 marks rhombomeres 3 and 5 of the hindbrain and *XlHbox*-6 is a spinal cord marker. N-CAM is a general neural marker whose expression is not restricted along the anteroposterior axis. The EF-1α control demonstrates that a comparable amount of RNA was assayed in each set. Note that expression of XAG-1 and anterior neural markers is stimulated by X-*bhh* treatment, whereas expression of posterior neural markers is suppressed.

FIGURE 20 demonstration of differential activities of N and C domains of hedgehog proteins. As in FIGURE 18 above, odd numbered lanes are negative control lanes, and positive numbered lanes show specific gene expression for the markers described above. The N domain of hedgehog is encoded in the construct called Xhh1208 (lane 8), and the C domain is encoded in the construct called Xhh1delta 27-208 (lane 10). The construct Xhh11-1270A (lane 12) is specifically mutated so that it is unable to undergo self-processing. The ability of the N and C domains to induce the genes described above is compared to control blastula cap explants (lane 4), entire embryos as a positive control (lane 2), blastula cap explants expressing a mutated hedgehog as a negative control (lane 14), blastula caps expressing the entire hedgehog 1 (lane 6), and blastula cap explants treated with an independent neural inducer, noggin (lane 16) (discovered by Richard Harland at University of California at Berkeley).

Examining the first panel for the cement gland marker XAG-1 clearly shows that intact hedgehog (lane 6) and the N domain (lane 8) and the processing defective hedgehog (lane 12) are much better than inducing the cement gland than is the C domain (lane 1). Examining the second panel demonstrates that the C domain (lane 10) is better at

inducing the pituitary gene XANF1B than is the N domain (lane 8). Since the N domain induces the XAG-1 marker better, described in point A above, the two results together clearly demonstrate that the N and C domains have distinguishable activities. Examination of the remaining panels shows that all described activities of the normal hedgehog (lane 6) can be defined in terms of the activities of the N and C domain.

Examining the third panel, for the forebrain gene otx-A, shows that both the N domain (lane 8) and C domain (lane 10) induce similar levels of this gene, but the processing defective hedgehog (lane 12) is better than either at inducing this gene.

Examining the fourth panel of this figure (NCAM), (as well as the FIGURE 18 panels EN-2, krox20, XIHbox6, and NCAM), shows that hedgehog does not induces these more posterior neural genes. Notably, noggin (lane 16) is able to induce pituitary gene and forebrain gene, but it also induces the general neural gene, NCAM, which hedgehog does not. This clearly shows that hedgehog is a distinct activity from the neural inducer noggin, and has a more restricted ability to induce neural genes.

Experiments in the Xenopus embryo were conducted by injecting full-length hedgehog RNA, and immunoprecipitating with a C-domain specific antibody, which proves that full length hedgehog does in fact get processed *in vivo* in vertebrates, consistent with the data shown in earlier Examples in *Drosophila*. Thus, the ideas for the utility of detecting hedgehog N and C domains is based on knowledge that such domains do appear through hedgehog processing in vertebrates. Moreover, the knowledge that hedgehog processing does occur *in vivo* naturally raised the question of whether the resulting N and C domains have independent activity.

The results in FIGURE 18 are novel insofar as they establish that the activity of hedgehog in inducing a pituitary gene, and an anterior brain gene, may be enhanced by the TGFβ family of growth factors. This enhancement likely applies to the N and C domains described in FIGURE 20, since the genes analyzed are the same. This enhancement is due

to *hh* synergizing with neural inducing factors which are themselves induced by TGF-β family members, including but not limited to such molecules as noggin and follistatin.

The data in FIGURE 20 makes several important points. First, the data show that the N and C domains have different though somewhat overlapping activities, and that the N and C activities added together account for all of the observed activity of the intact hedgehog protein. Thus, any clinical or diagnostic uses of hedgehog might be improved by use of the N or C domain, as one generally wishes to use the smallest protein which has an activity for clinical work, as it is less likely to evoke adverse immune responses, or other adverse side effects. Second, the data show that the C domain is better than the N domain in inducing pituitary gene expression and, since it has less induction of cement gland genes that intact hedgehog, or N domain, it suggests that the C domain might be useful in clinical situations where one wishes to enhance the development or expression of the pituitary as specifically as possible. As the pituitary is the source of a number of hormones, any treatment for enhancing pituitary cell growth and activity would ideally have as few side effects as possible, and the C domain is thus a viable candidate for therapies with enhanced pituitary cell growth and function in mind. Third, relating to studies regarding noggin, FIGURE 20 shows clearly that while both hedgehog and noggin can induce pituitary gene expression, hedgehog is more specific, since hedgehog does not induce the general neural marker NCAM, whereas noggin induces NCAM as well as pituitary. Fourth, the hedgehog which was mutated to prevent processing (lane 12) is as active as full-length and wild-type hedgehog (lane 6) in inducing pituitary gene expression, but the processing defective hedgehog is better at inducing the forebrain marker otx-A. Thus, for some clinical applications of hedgehog in inducing specific cell types, it is possible that the processing-defective hedgehog will be superior compared to normal hedgehog.

FIGURE 21 shows ΔN-C interferes with X-*bhh* and N activity in animal cap explants. Embryos were injected with various RNAs, animal cap explants were cultured until sibling embryos reached tailbud (stage 25), at which time RT-PCR was used to analyze

the expression of the cement gland marker XAG-1 and the control RNA, EF-1α. Lanes 1, 2: control animal caps from uninjected embryos. Lanes 3, 4: animal caps from embryos injected with both X-*bhh* and prolactin RNAs. Lanes 5, 6: animal caps from embryos injected with box X-*bhh* and ΔN-C. Lanes 7, 8: animal caps from embryos injected with both N and prolacting RNAs. Lanes 9, 10: animal caps from embryos injected with both N and ΔN-C. The N and X-*bhh* experiments were conducted independently and thus absolute levels in lanes 3-6 should not be compared to those in lanes 7-10. Note that the induction of XAG-1 expression by X-*bhh* or N is reduced by co-injection of ΔN-C.

An internal deletion of *X-bhh* (ΔN-C) blocked the activity of *X-bhh* and N in explants and reduced dorsoanterior structures in embryos. As elevated *hh* activity increases the expression of anterior neural genes, and as ΔN-C reduces dorsoanterior structures, these complementary data support a role for *hh* in neural induction and anteroposterior patterning.

ΔN-C deletes amino acids 28-194 of X-bhh. The primary translation product is predicted to undergo signal sequence cleavage removing amino acids 1-23, and to undergo autoproteolysis. Based on the cleavage site in Drosophila *hh* (Porter, *et al.*, *Nature*, 374:363, 1995) autoproteolysis would generate a C domain of X-bhh amino acids 198-409, as well as a predicted seven amino acid polypeptide, representing amino acids 24-27, and 195-197 (Lai, *et al.*, *Development* 121:2349, 1995). Analysis of the effect of ΔN-C on neural markers was by standard methods including Northern blot analysis and in situ hybridization (Lai, *et al.*, *supra*, incorporated herein by reference).

Although ΔN-C does not induce the cement gland marker XAG-1, it decreases the expression of anterior ectodermal and neural markers in activin-treated animal caps. Thus, ΔN-C has the capacity to affect neural patterning. ΔN-C also promotes an increase in posterior neural markers in activin-treated animal caps. Mixing ΔN-C with N or full length X-*bhh* at a 1:1 ratio led to a dramatic inhibition of the induction of cement gland in animal cap assays, supporting the hypothesis that ΔN-C interfered with X-*hh*.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only and are not intended to limit the scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: The Johns Hopkins University School of Medicine, et al.

(ii) TITLE OF INVENTION: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 4225 Executive Square, Suite 1400
(C) CITY: La Jolla
(D) STATE: CA
(E) COUNTRY: U.S.A.
(F) ZIP: 92037

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US95
(B) FILING DATE: 01-DEC-1995
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Haile, Lisa A.
(B) REGISTRATION NUMBER: 38,347
(C) REFERENCE/DOCKET NUMBER: 07265/043WO1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619/678-5070
(B) TELEFAX: 619/678-5099

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 144 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTG AAA CTG CGG GTG ACC GAG CCC TGG GAC GAA GAT GGC CAC CAC TCA
48
Val Lys Leu Arg Val Thr Glu Pro Trp Asp Glu Asp Gly His His Ser
1 5 10 15

CAG GAG TCT CTG CAC TAC GAG GGC CGC GCA GTG GAC ATC ACC ACG TCT
96
Gln Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser
20 25 30

GAC CGC GAC CGC AGC AAG TAC GGC ATG CTG GCC CGC CTG GCG GTG G
142
Asp Arg Asp Arg Ser Lys Tyr Gly Met Leu Ala Arg Leu Ala Val
35 40 45

AG
144

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 144 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTG AAG CTG CGG GTG ACC GAG GGC TGG GAC GAG GAC GGC CAC CAC TCA
48
Val Lys Leu Arg Val Thr Glu Gly Trp Asp Glu Asp Gly His His Ser
50 55 60

GAG GAG TCC CTG CAT TAT GAG GGC CGC GCG GTG GAC ATC ACC ACA TCA
96
Glu Glu Ser Leu His Tyr Glu Gly Arg Ala Val Asp Ile Thr Thr Ser
65 70 75

GAC CGC GAC CGC AAT AAG TAT GGA CTG CTG GCG CGC TTG GCA GTG G
142
Asp Arg Asp Arg Asn Lys Tyr Gly Leu Leu Ala Arg Leu Ala Val
80 85 90

AG
144

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ile Ser Ser His Val His Gly Cys Phe Thr Pro Glu Ser Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ser Ile Ser His Met His Gly Cys Phe Thr Pro Glu Ser Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Val Ala Ala Lys Ser Asp Gly Cys Phe Pro Gly Ser Ala Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Gly Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Ala Gly Ala Arg Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Ala Ala Lys Thr Gly Gly Cys Phe Pro Ala Gly Ala Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Ala Ala Lys Thr Gly Gly Cys Phe Pro Gly Glu Ala Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Gly Val Arg Ser Gly Gly Cys Phe Pro Gly Thr Ala Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Leu Ala Val Arg Ala Gly Gly Cys Phe Pro Gly Asn Ala Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

His Gly His Gly Cys Phe Thr Pro
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

His Gly His Gly Cys Phe Thr Pro
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Lys Ser Gly Gly Cys Phe Pro Gly
1 5

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 416 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Asp Val Arg Leu His Leu Lys Gln Phe Ala Leu Leu Cys Phe Ile
1 5 10 15

Ser Leu Leu Leu Thr Pro Cys Gly Leu Ala Cys Gly Pro Gly Arg Gly
20 25 30

Tyr Gly Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
35 40 45

Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Lys
50 55 60

Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Ile
65 70 75 80

Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Asn
85 90 95

Ala Asp Arg Leu Met Thr Lys Arg Cys Lys Asp Lys Leu Asn Ser Leu
100 105 110

Ala Ile Ser Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr
115 120 125

Glu Gly Trp Asp Glu Asp Gly His His Leu Glu Glu Ser Leu His Tyr
130 135 140

Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys
145 150 155 160

Tyr Gly Met Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val
165 170 175

Tyr Tyr Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn
180 185 190

Ser Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Gly Thr Val
195 200 205

Thr Leu Gly Asp Gly Thr Arg Lys Pro Ile Lys Asp Leu Lys Val Gly
210 215 220

Asp Arg Val Leu Ala Ala Asp Glu Lys Gly Asn Val Leu Ile Ser Asp
225 230 235 240

Phe Ile Met Phe Ile Asp His Asp Pro Thr Thr Arg Arg Gln Phe Ile
245 250 255

Val Ile Glu Thr Ser Glu Pro Phe Thr Lys Leu Thr Leu Thr Ala Ala
260 265 270

His Leu Val Phe Val Gly Asn Ser Ser Ala Ala Ser Gly Ile Thr Ala
275 280 285

Thr Phe Ala Ser Asn Val Lys Pro Gly Asp Thr Val Leu Val Trp Glu
290 295 300

Asp Thr Cys Glu Ser Leu Lys Ser Val Thr Val Lys Arg Ile Tyr Thr
305 310 315 320

Glu Glu His Glu Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile
325 330 335

Ile Val Asp Gln Val Leu Ala Ser Cys Tyr Ala Val Ile Glu Asn His
340 345 350

Lys Trp Ala His Trp Ala Phe Ala Pro Val Arg Leu Cys His Lys Leu
355 360 365

Met Thr Trp Leu Phe Pro Ala Arg Glu Ser Asn Val Asn Phe Gln Glu
370 375 380

Asp Gly Ile His Trp Tyr Ser Asn Met Leu Phe His Ile Gly Ser Trp
385 390 395 400

Leu Leu Asp Arg Asp Ser Phe His Pro Leu Gly Ile Leu His Leu Ser
405 410 415

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Arg Leu Leu Thr Arg Val Leu Leu Val Ser Leu Leu Thr Leu Ser
1 5 10 15

Leu Val Val Ser Gly Leu Ala Cys Gly Pro Gly Arg Gly Tyr Gly Arg
20 25 30

Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe Ile
35 40 45

Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu Gly
50 55 60

Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn Tyr
65 70 75 80

Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp Arg
85 90 95

Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ser Leu Ala Ile Ser
100 105 110

Val Met Asn His Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly Trp
115 120 125

Asp Glu Asp Gly His His Phe Glu Glu Ser Leu His Tyr Glu Gly Arg
130 135 140

Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Lys Ser Lys Tyr Gly Thr
145 150 155 160

Leu Ser Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr Glu
165 170 175

Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val Ala
180 185 190

Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Leu Val Ser Leu Gln
195 200 205

Asp Gly Gly Gln Lys Ala Val Lys Asp Leu Asn Pro Gly Asp Lys Val
210 215 220

Leu Ala Ala Asp Ser Ala Gly Asn Leu Val Phe Ser Asp Phe Ile Met
225 230 235 240

Phe Thr Asp Arg Asp Ser Thr Thr Arg Arg Val Phe Tyr Val Ile Glu
245 250 255

Thr Gln Glu Pro Val Glu Lys Ile Thr Leu Thr Ala Ala His Leu Leu
260 265 270

Phe Val Leu Asp Asn Ser Thr Glu Asp Leu His Thr Met Thr Ala Ala
275 280 285

Tyr Ala Ser Ser Val Arg Ala Gly Gln Lys Val Met Val Val Asp Asp
290 295 300

Ser Gly Gln Leu Lys Ser Val Ile Val Gln Arg Ile Tyr Thr Glu Glu
305 310 315 320

Gln Arg Gly Ser Phe Ala Pro Val Thr Ala His Gly Thr Ile Val Val
325 330 335

Asp Arg Ile Leu Ala Ser Cys Tyr Ala Val Ile Glu Asp Gln Gly Leu
340 345 350

Ala His Leu Ala Phe Ala Pro Ala Arg Leu Tyr Tyr Tyr Val Ser Ser
355 360 365

Phe Leu Phe Pro Gln Asn Ser Ser Ser Arg Ser Asn Ala Thr Leu Gln
370 375 380

Gln Glu Gly Val His Trp Tyr Ser Arg Leu Leu Tyr Gln Met Gly Thr
385 390 395 400

Trp Leu Leu Asp Ser Asn Met Leu His Pro Leu Gly Met Ser Val Asn
405 410 415

Ser Ser

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 425 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: not relevant
(D) TOPOLOGY: both

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Val Glu Met Leu Leu Leu Thr Arg Ile Leu Leu Val Gly Phe Ile
1 5 10 15

Cys Ala Leu Leu Val Ser Ser Gly Leu Thr Cys Gly Pro Gly Arg Gly
20 25 30

Ile Gly His Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys
35 40 45

Gln Phe Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg
50 55 60

Tyr Glu Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Ile
65 70 75 80

Pro Asn Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly
85 90 95

Ala Asp Arg Leu Met Thr Cys Arg Cys Lys Asp Lys Leu Asn Ala Leu
100 105 110

Ala Ile Ser Val Met Asn Cys Trp Pro Gly Val Met Leu Arg Val Thr
115 120 125

Glu Gly Trp Asp Glu Asp Gly His His Ser Lys Glu Ser Leu His Tyr
130 135 140

Glu Gly Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys
145 150 155 160

Tyr Gly Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val
165 170 175

Tyr Tyr Glu Ser Lys Ala His Ile Cys Ser Val Lys Ala Glu Asn Ser
180 185 190

Val Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His
195 200 205

Leu Glu His Gly Gly Thr Lys Leu Val Lys Asp Leu Ser His Gly Asp
210 215 220

Arg Val Leu Ala Ala Asp Ala Asp Gly Arg Leu Leu Val Ser Asp Phe
225 230 235 240

Leu Leu Thr Phe Leu Asp Arg Met Asp Ser Ser Arg Lys Leu Phe Tyr
245 250 255

Val Ile Glu Thr Arg Gln Pro Arg Ala Arg Leu Leu Leu Thr Ala Ala
260 265 270

His Leu Leu Phe Val Ala Pro Gln His Asn Gln Ser Glu Ala Thr Gly
275 280 285

Ser Thr Ser Gly Gln Ala Leu Phe Ala Ser Asn Val Lys Pro Gly Gln
290 295 300

Pro Val Val Val Leu Gly Glu Gly Gly Gln Gln Leu Leu Pro Ala Ser
305 310 315 320

Val His Ser Val Ser Leu Arg Glu Glu Ala Ser Gly Ala Tyr Ala Pro
325 330 335

Thr Thr Ala Cys Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys
340 345 350

Tyr Ala Val Ile Glu Glu His Ser Trp Ala His Ala Ala Phe Ala Pro
355 360 365

His Arg Leu Ala Gln Gly Leu Leu Ala Ala Leu Cys Pro Asp Gly Ala
370 375 380

Ile Pro Thr Ala Ala Thr Thr Thr Thr Gly Ile His Trp Tyr Ser Arg
385 390 395 400

Leu Leu Tyr Arg Ile Gly Ser Trp Val Leu Asp Gly Asp Ala Leu His
405 410 415

Pro Leu Gly Met Val Ala Pro Ala Ser
420 425

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 437 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Leu Leu Leu Leu Ala Arg Cys Phe Leu Val Ile Leu Ala Ser Ser
1 5 10 15

Leu Leu Val Cys Pro Gly Leu Ala Cys Gly Pro Gly Arg Gly Phe Gly
20 25 30

Lys Arg Arg His Pro Lys Lys Leu Thr Pro Leu Ala Tyr Lys Gln Phe
35 40 45

Ile Pro Asn Val Ala Glu Lys Thr Leu Gly Ala Ser Gly Arg Tyr Glu
50 55 60

Gly Lys Ile Thr Arg Asn Ser Glu Arg Phe Lys Glu Leu Thr Pro Asn
65 70 75 80

Tyr Asn Pro Asp Ile Ile Phe Lys Asp Glu Glu Asn Thr Gly Ala Asp
85 90 95

Arg Leu Met Thr Gln Arg Cys Lys Asp Lys Leu Asn Ala Leu Ala Ile
100 105 110

Ser Val Met Asn Gln Trp Pro Gly Val Lys Leu Arg Val Thr Glu Gly
115 120 125

Trp Asp Glu Asp Gly His His Ser Glu Glu Ser Leu His Tyr Glu Gly
130 135 140

Arg Ala Val Asp Ile Thr Thr Ser Asp Arg Asp Arg Ser Lys Tyr Gly
145 150 155 160

Met Leu Ala Arg Leu Ala Val Glu Ala Gly Phe Asp Trp Val Tyr Tyr
165 170 175

Glu Ser Lys Ala His Ile His Cys Ser Val Lys Ala Glu Asn Ser Val
180 185 190

Ala Ala Lys Ser Gly Gly Cys Phe Pro Gly Ser Ala Thr Val His Leu
195 200 205

Glu Gln Gly Gly Thr Lys Leu Val Lys Asp Leu Arg Pro Gly Asp Arg
210 215 220

Val Leu Ala Ala Asp Asp Gln Gly Arg Leu Leu Tyr Ser Asp Phe Leu
225 230 235 240

Thr Phe Leu Asp Arg Asp Glu Gly Ala Lys Lys Val Phe Tyr Val Ile
245 250 255

Glu Thr Leu Glu Pro Arg Glu Arg Leu Leu Leu Thr Ala Ala His Leu
260 265 270

Leu Phe Val Ala Pro His Asn Asp Ser Gly Pro Thr Pro Gly Pro Ser
275 280 285

Ala Leu Phe Ala Ser Arg Val Arg Pro Gly Gln Arg Val Tyr Val Val
290 295 300

Ala Glu Arg Gly Gly Asp Arg Arg Leu Leu Pro Ala Ala Val His Ser
305 310 315 320

Val Thr Leu Arg Glu Glu Glu Ala Gly Ala Tyr Ala Pro Leu Thr Ala
325 330 335

His Gly Thr Ile Leu Ile Asn Arg Val Leu Ala Ser Cys Tyr Ala Val
340 345 350

Ile Glu Glu His Ser Trp Ala His Arg Ala Phe Ala Pro Phe Arg Leu
355 360 365

Ala His Ala Leu Leu Ala Ala Leu Ala Pro Ala Arg Thr Asp Gly Gly
370 375 380

Gly Gly Gly Ser Ile Pro Ala Ala Gln Ser Ala Thr Glu Ala Arg Gly
385 390 395 400

Ala Glu Pro Thr Ala Gly Ile His Trp Tyr Ser Gln Leu Leu Tyr His
405 410 415

Ile Gly Thr Trp Leu Leu Asp Ser Glu Thr Met His Pro Leu Gly Met
420 425 430

Ala Val Lys Ser Ser
435

CLAIMS

1. A substantially pure polypeptide characterized by having an amino acid sequence derived from amino terminal amino acids of a hedgehog protein and having at its carboxy terminus, a G↓CF cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide.

2. An isolated polynucleotide sequence encoding a polypeptide having an amino acid sequence of claim 1.

3. The isolated polynucleotide sequence of claim 2, encoding a polypeptide having an amino acid sequence of the polypeptide of claim 1 and having at least one epitope for an antibody immunoreactive with the polypeptide of claim 1.

4. A recombinant expression vector which contains the polynucleotide of claim 2.

5. A host cell which contains the expression vector of claim 4.

6. An antibody which binds to the polypeptide of claim 1 and which binds with immunoreactive fragments of the polypeptide of claim 1.

7. The antibody of claim 6, wherein the antibody is polyclonal.

8. The antibody of claim 6, wherein the antibody is monoclonal.

9. A substantially pure polypeptide characterized by having an amino acid sequence derived from carboxy terminal amino acids of a hedgehog protein and having at its amino terminus, a G↓CF cleavage site specifically recognized by a proteolytic activity of the carboxy terminal fragment of the native hedgehog polypeptide.

10. An isolated polynucleotide sequence encoding a polypeptide having an amino acid sequence of claim 9.

11. The isolated polynucleotide sequence of claim 10, encoding a polypeptide having an amino acid sequence of the polypeptide of claim 9 and having at least one epitope for an antibody immunoreactive with the polypeptide of claim 9.

12. A recombinant expression vector which contains the polynucleotide of claim 10.

13. A host cell which contains the expression vector of claim 12.

14. An antibody which binds to the polypeptide of claim 9 and which binds with immunoreactive fragments of the polypeptide of claim 9.

15. The antibody of claim 14, wherein the antibody is polyclonal.

16. The antibody of claim 14, wherein the antibody is monoclonal.

17. A method for modulating proliferation or differentiation of neuronal cells, comprising contacting the cells with a hedgehog polypeptide.

18. The method of claim 17, wherein the hedgehog polypeptide is the polypeptide of claim 1.

19. The method of claim 17, wherein the hedgehog polypeptide is the polypeptide of claim 9.

20. The method of claim 17, wherein the modulation is induction of proliferation or differentiation.

21. The method of claim 17, wherein the neuronal cells are substantially derived from floor plate neuronal cells.

22. The method of claim 21, wherein the neuronal cells are motor neurons.

23. The method of claim 17, wherein the neuronal cells are vertebrate neuronal cells.

24. The method of claim 23, wherein the vertebrate is a human.

25. The method of claim 17, further comprising contacting the cell with a member of the TGF-β family.

26. The method of claim 25, wherein the TGF-β is activin.

27. The method of claim 17, further comprising contacting the cell with a neural inducing agent.

28. The method of claim 27, wherein the neural inducing agent is selected from the group consisting of noggin, follistatin, and a neurotrophin.

29. An autoproteolytic fusion protein comprising a first polypeptide including the proteolytic domain of the polypeptide of claim 9, a cleavage site recognized by the first polypeptide, and a second polypeptide.

30. A polynucleotide encoding the fusion protein of claim 29.

31. A method for producing an autoproteolytic fusion protein comprising operably linking a first polynucleotide encoding a first polypeptide including the proteolytic domain of the polypeptide of claim 9 and the cleavage site recognized by the proteolytic domain, and a second polynucleotide encoding a second polypeptide.

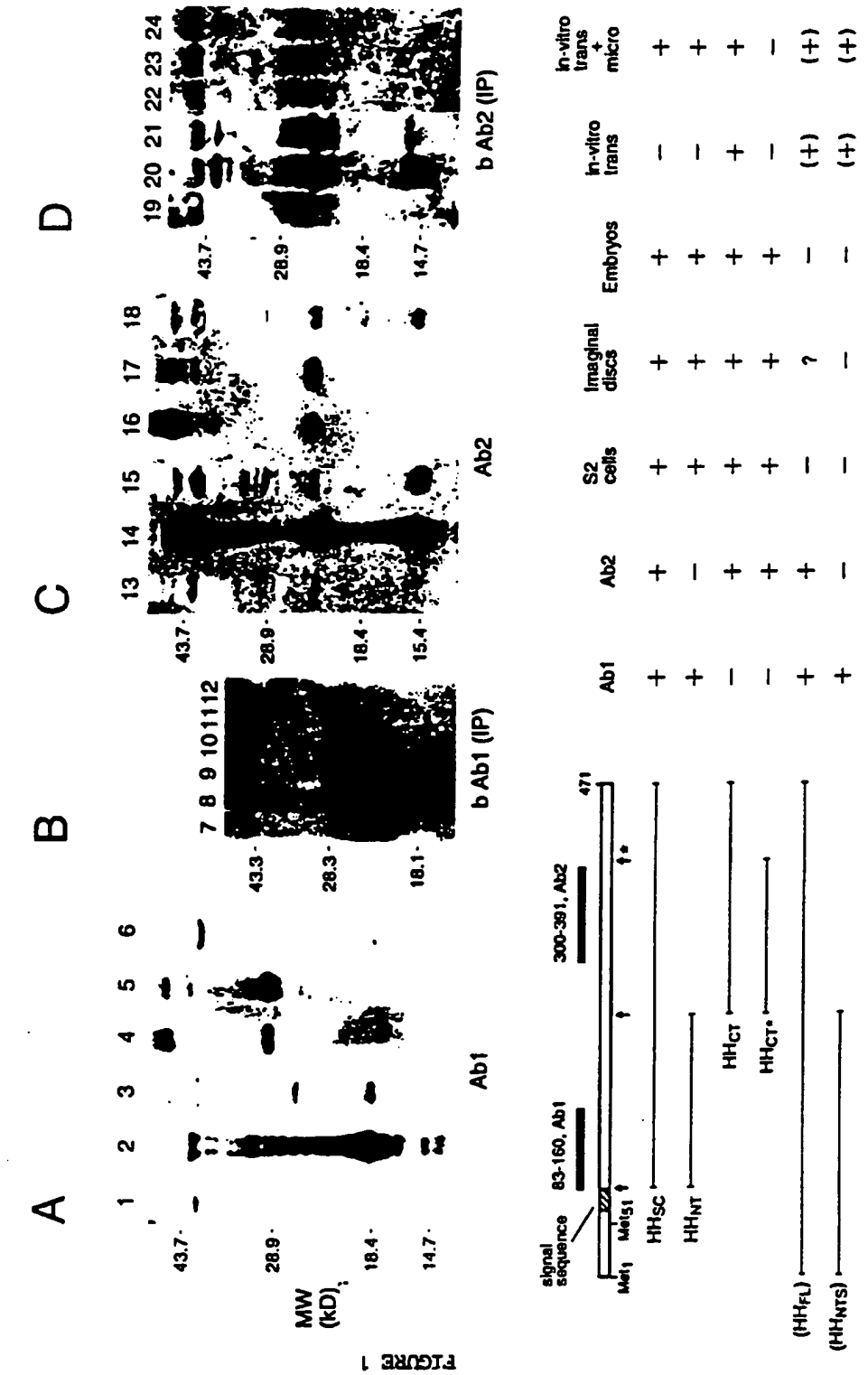

A
B
C
D
MW (kD)
Ab1
b Ab1 (IP)
Ab2
b Ab2 (IP)
signal sequence
83-160, Ab1
300-391, Ab2
471
Met1
Met51
HH5C
HHNT
HHCT
HHCT*
(HHFL)
(HHNTS)
Ab1
Ab2
S2 cells
Imaginal discs
Embryos
In-vitro trans
In-vitro trans + micro

FIGURE 1

FIGURE 1E

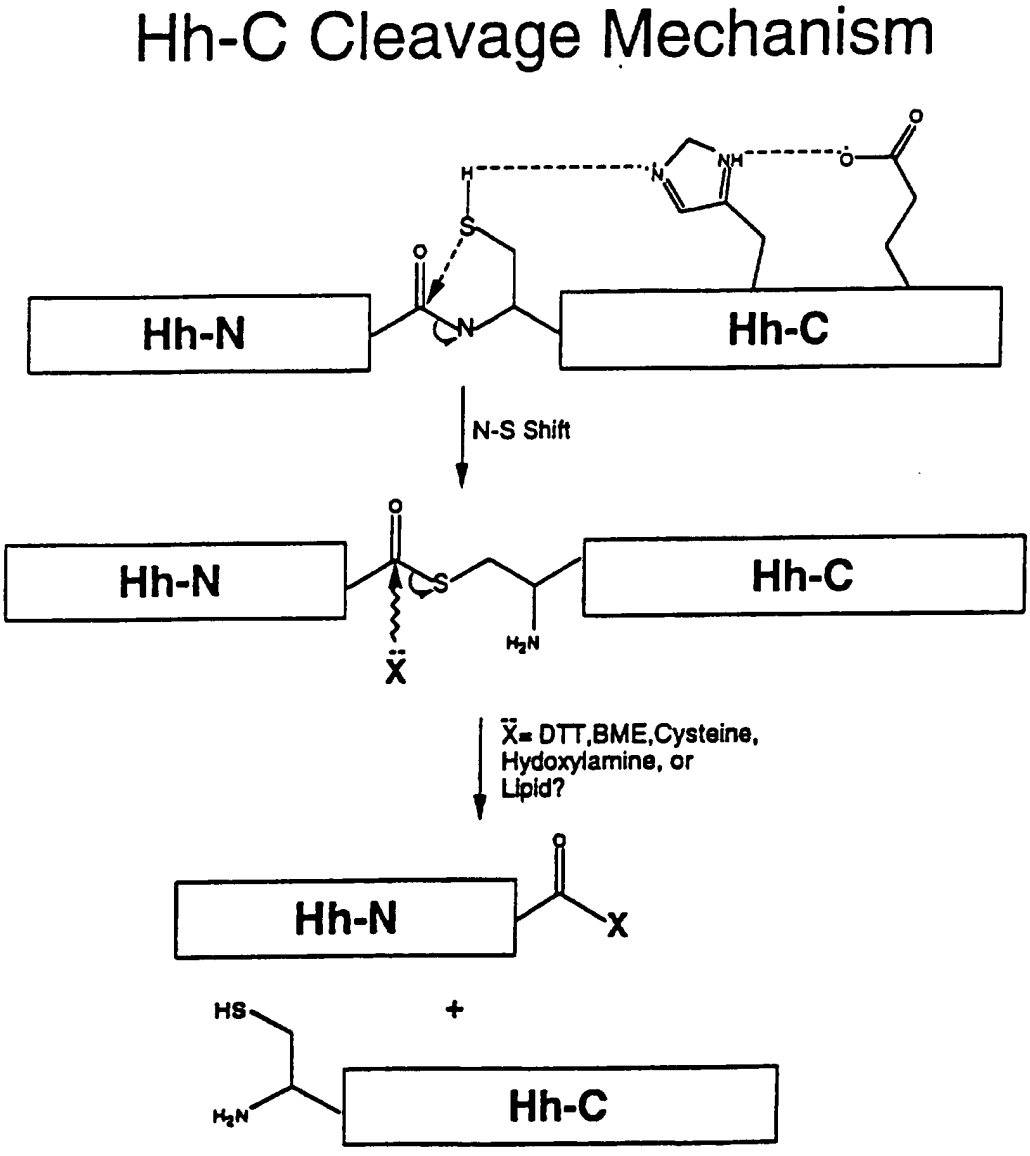

Hh-C Cleavage Mechanism
Hh-N
Hh-C
N-S Shift
Hh-N
Hh-C
X= DTT,BME,Cysteine, Hydoxylamine, or Lipid?
Hh-N
X
HS
+
Hh-C

FIGURE
2

		1	2	3	4	5	6	7
A	D. mel. hh	L	T	V	T	P	A	H
	D. hydei hh	L	T	V	T	P	A	H
	C-Shh	L	L	L	T	A	A	H
	M-Shh/Hhg-1	L	L	L	T	A	A	H
	R vhh-1	L	L	L	T	A	A	H
	Z-Shh/Zf vhh-1	I	T	L	T	A	A	H
	twhh	L	T	L	T	A	A	H
	M-Dhh	L	L	L	T	P	W	H
	M-Ihh	L	A	L	T	P	A	H
B	CHT	W	V	V	T	A	A	H
	TRP	W	V	V	S	A	A	H
	ELA	W	V	M	T	A	A	H
	UKH	W	V	I	S	A	T	H
	C1R	W	I	L	T	A	A	H
	C1S	W	V	L	T	A	A	H
	MCP	F	V	L	T	A	A	H
	FAX	Y	V	L	T	A	A	H
	TPA	W	I	L	S	A	A	H

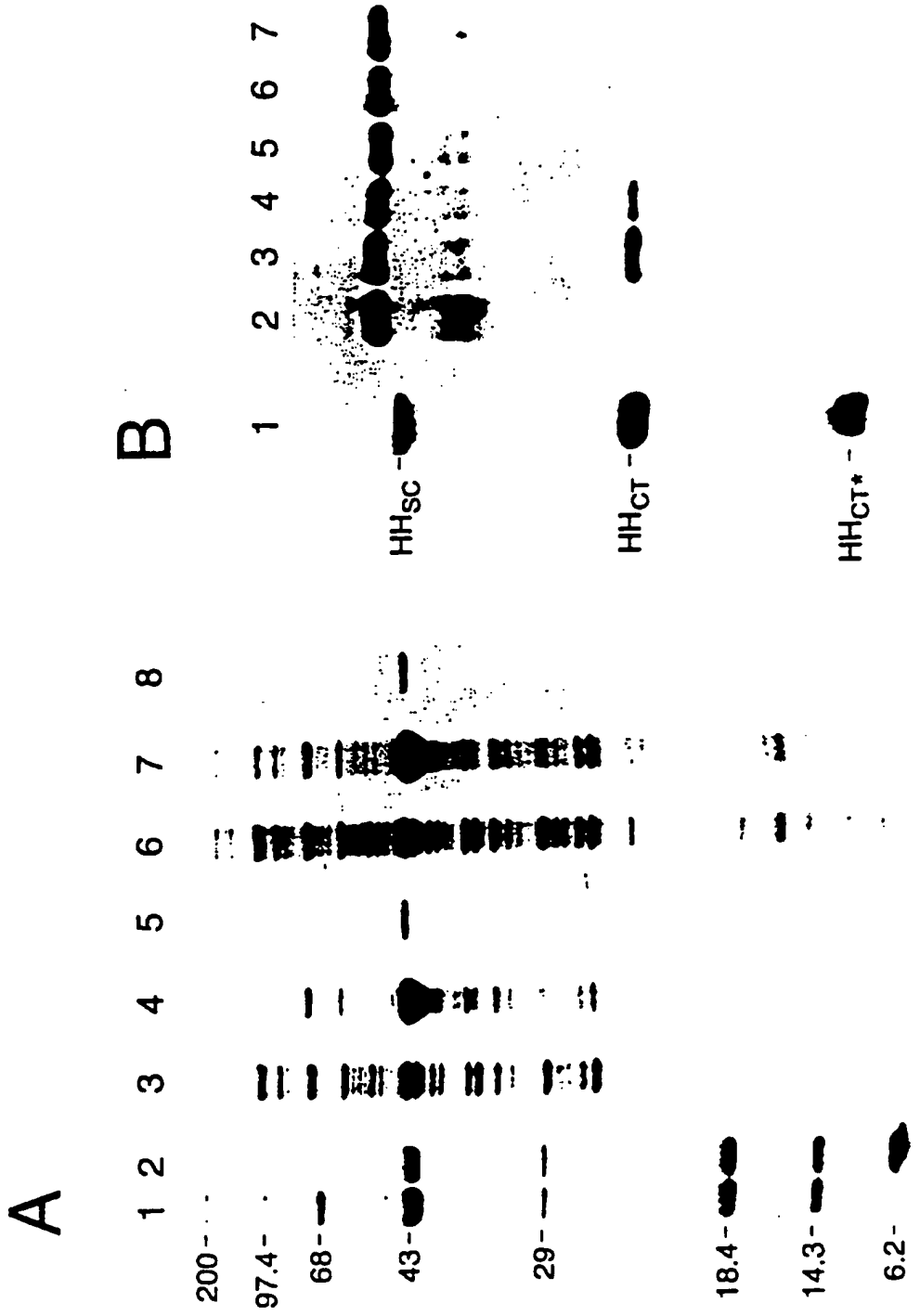
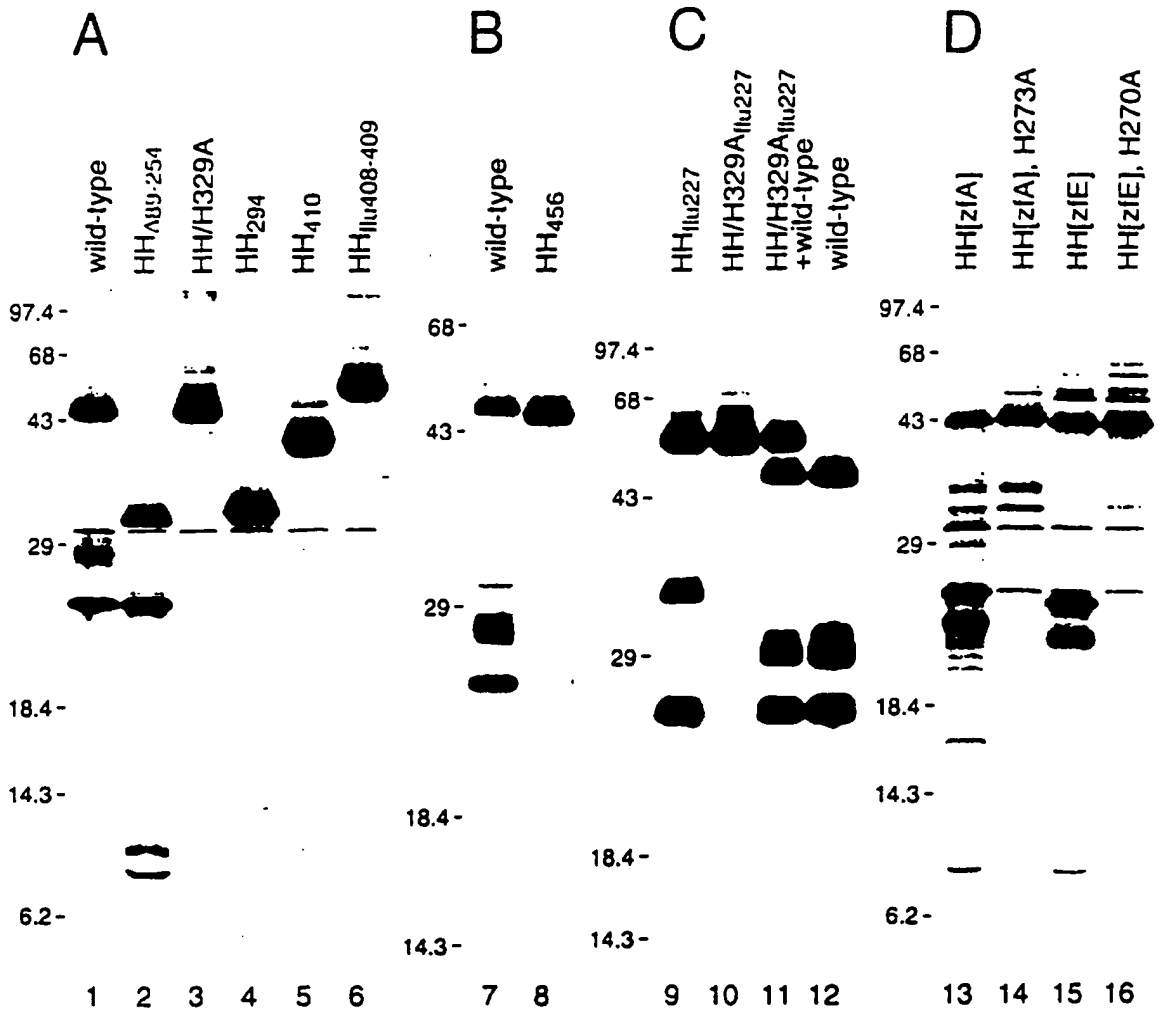
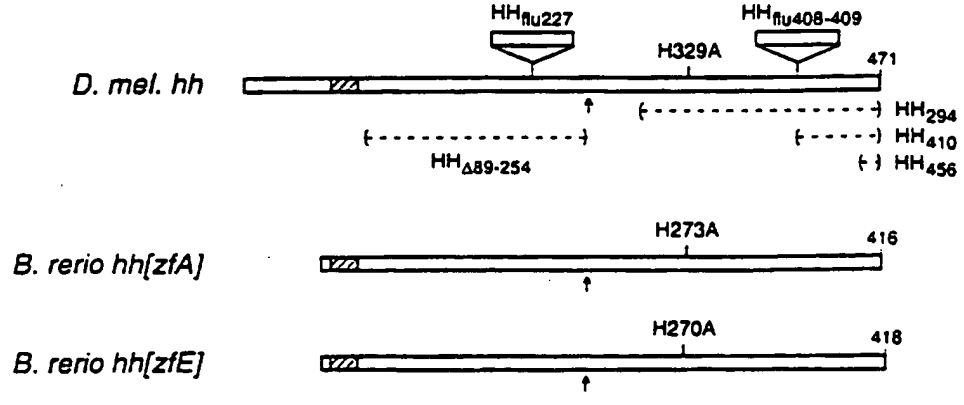


FIGURE 3

FIGURE 4


A
B
C
D
wild-type
HH$_{\Delta 89-254}$
HH/H329A
HH$_{294}$
HH$_{410}$
HH$_{flu408-409}$
wild-type
HH$_{456}$
HH$_{flu227}$
HH/H329A$_{flu227}$
HH/H329A$_{flu227}$ +wild-type
wild-type
HH[zfA]
HH[zfA], H273A
HH[zfE]
HH[zfE], H270A
97.4
68
43
29
18.4
14.3
6.2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16


D. mel. hh
HH$_{flu227}$
HH$_{flu408-409}$
H329A
471
HH$_{294}$
HH$_{410}$
HH$_{456}$
HH$_{\Delta 89-254}$
B. rerio hh[zfA]
H273A
416
B. rerio hh[zfE]
H270A
418

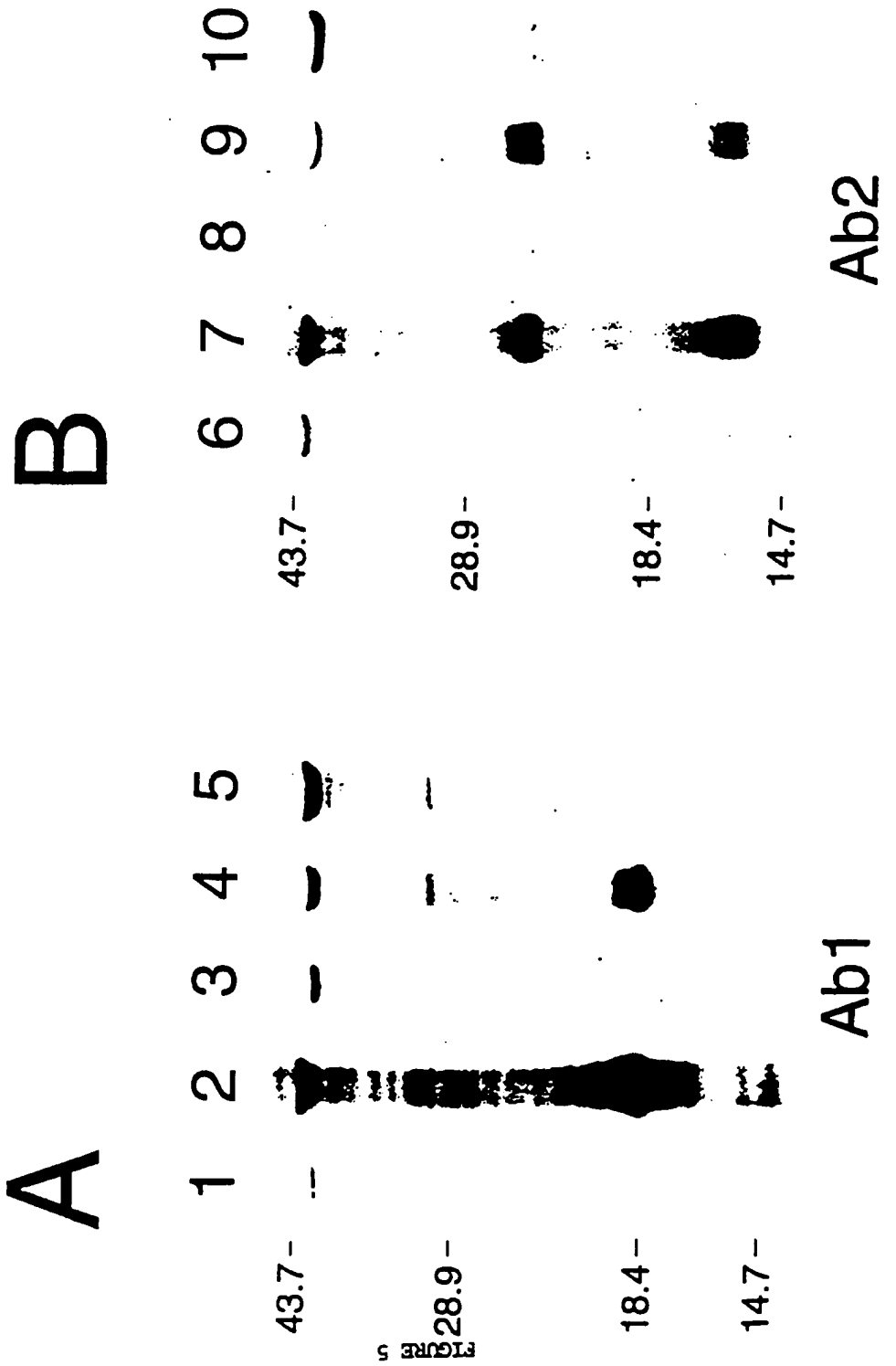
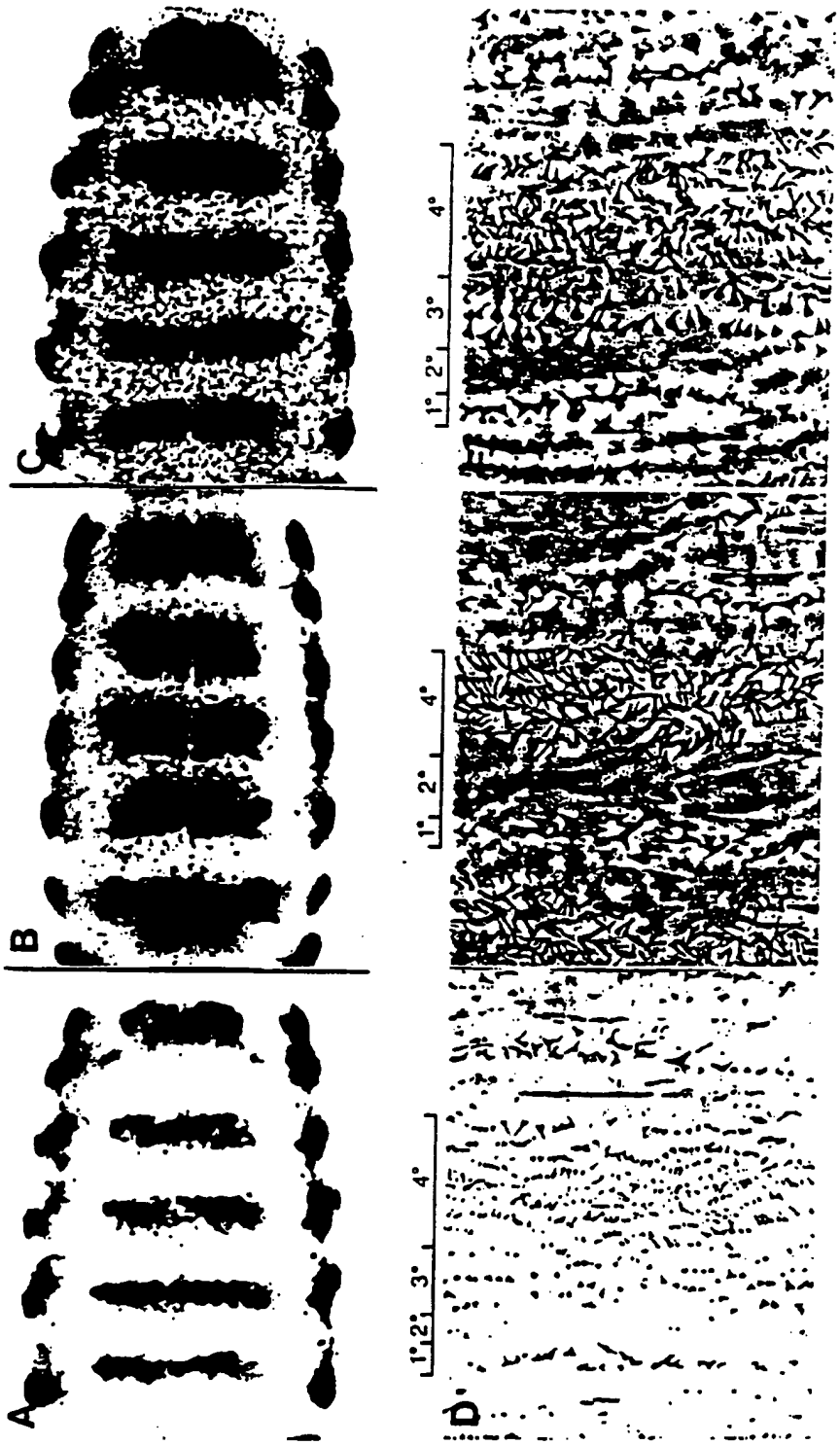
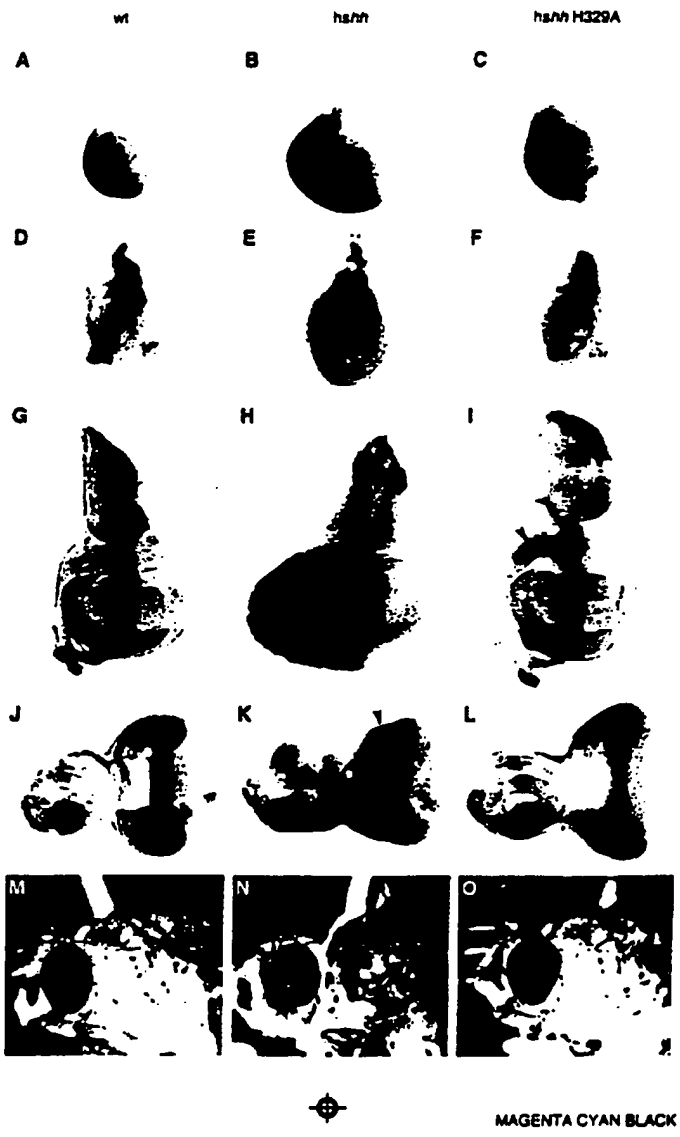


FIGURE 5

FIGURE 6



Lee Fig. 7



MAGENTA CYAN BLACK

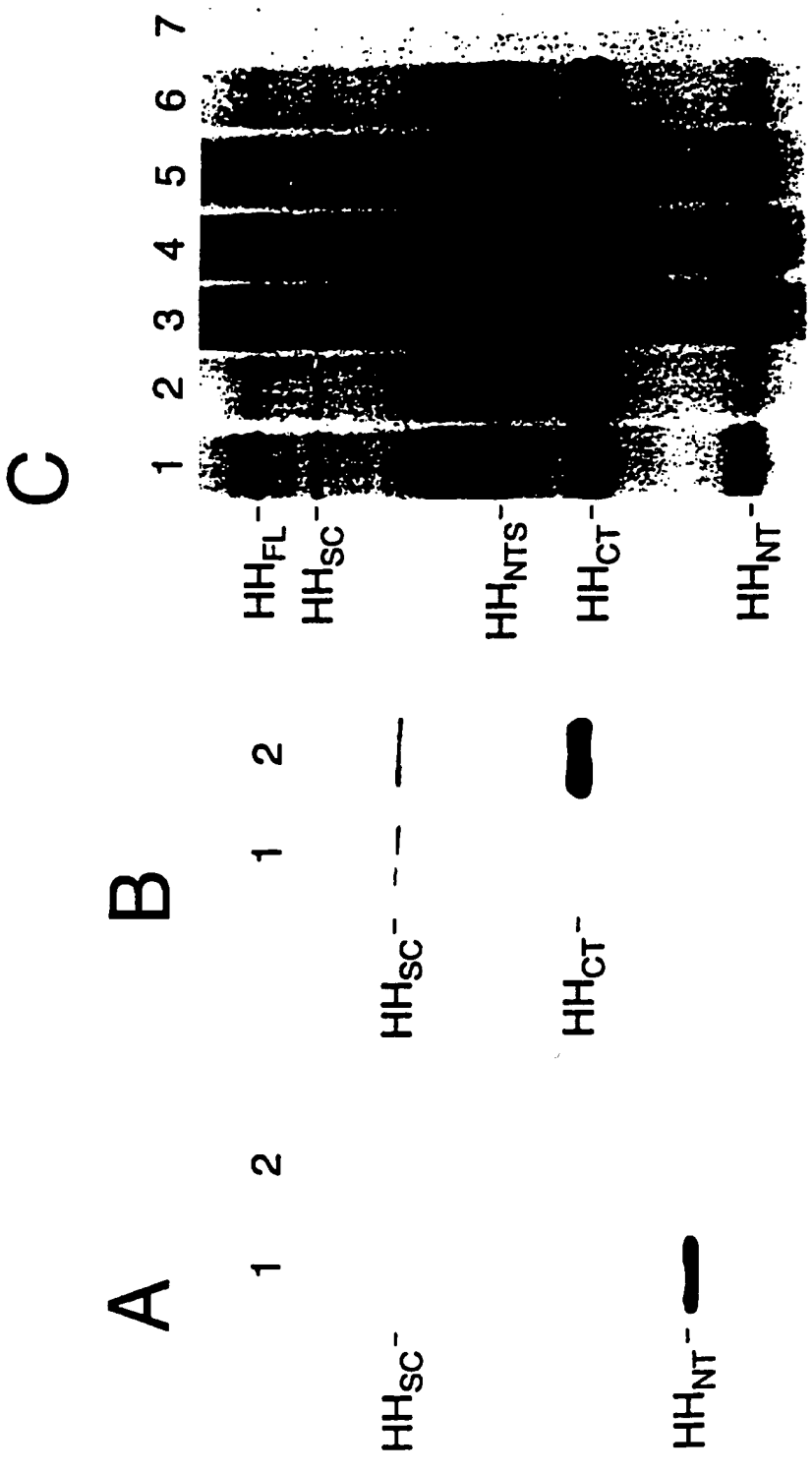
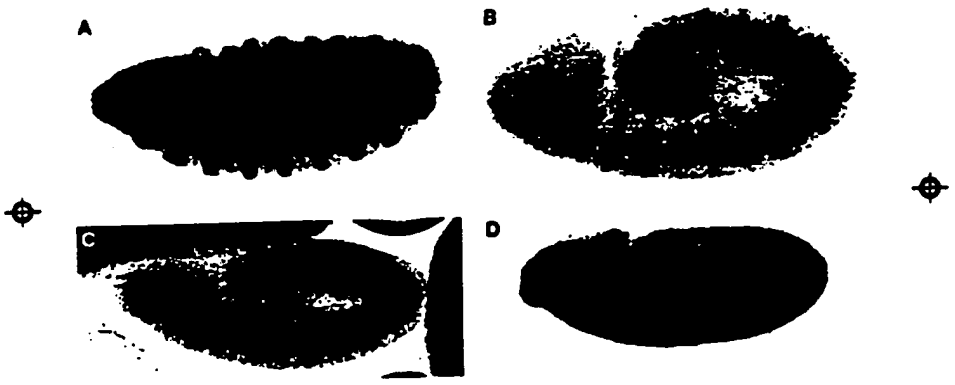

C
1 2 3 4 5 6 7
HH_FL
HH_SC
HH_NTS
HH_CT
HH_NT
B
1 2
HH_SC
HH_CT
A
1 2
HH_SC
HH_NT

FIGURE 8

Lee Fig. 9



MAGENTA CYAN BLACK

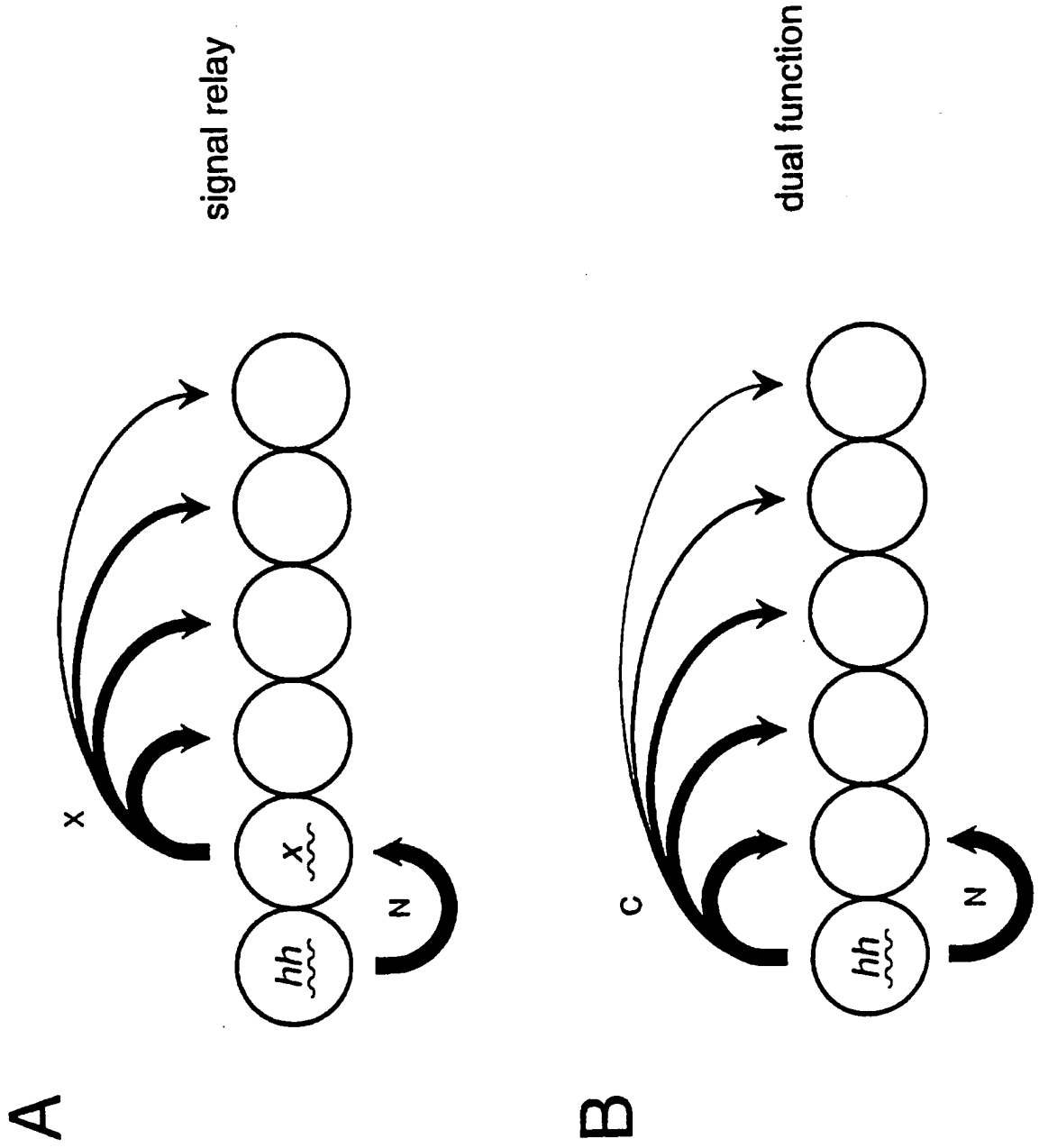


FIGURE 10

FIGURE 10

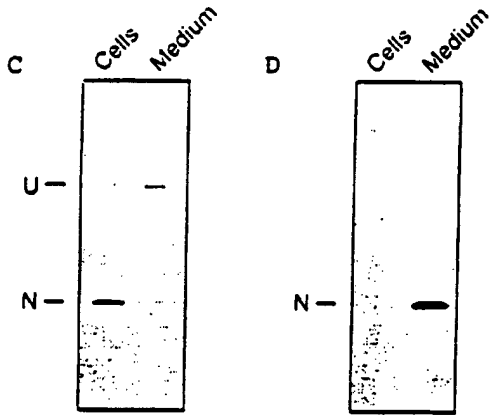


FIGURE 11A

human hh B -> 1-phase Translation Sonic HHG1

DNA sequence 144 b.p. GTGAAACTGCGC ... CTGGCGGTGGAG linear

```
1/1                                                 31/11
GTG AAA CTG CGC GTC ACC GAG GGC TGG GAC GAA GAT GGC CAC CAC TCA GAG GAG TCT CTG
V   K   L   R   V   T   E   G   W   D   E   D   G   H   H   S   E   E   S   L
61/21                                               91/31
GAC TAC GAG GGC CGC GCA GTG GAC ATC ACC ACG TCT GAC CGC GAC CGC AGC AAG TAC GGC
E   Y   E   G   R   A   V   D   I   T   T   S   D   R   D   R   S   K   Y   G
121/41
ATG CTG GCC CGC CTC GCG GTG GAG
M   L   A   R   L   A   V   E
```

FIGURE 11B

human hh A -> I-phase Translation HHG 2 (probably Desert Indian)

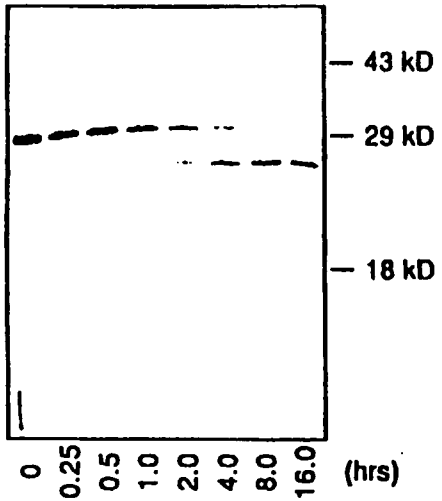
DNA sequence 144 b.p. GTGAAGCTGCGG ... TTGGCAGTGCAG linear

```
1/1                                                 31/11
GTG AAG CTG CGG GTG ACC GAG GGC TGG GAC GAG GAC GGC CAC CAC TTA GAG GAG TCC CTG
V   K   L   R   V   T   E   G   W   D   E   D   G   H   H   S   E   E   S   L
61/21                                               91/31
CAT TAT GAG GGC CGC GCG GTG GAC ATC ACC ACA TCA GAC CGC GAC CGC AAT AAG TAT GGA
H   Y   E   G   R   A   V   D   I   T   T   S   D   R   D   R   N   K   Y   G
121/41
CTG CTG GCG CGC TTG GCA GTG GAG
L   L   A   R   L   A   V   E
```

FIGURE 12

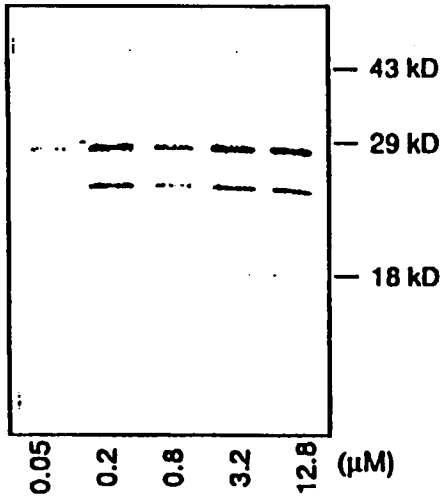
A

Autoproteolysis Time Course



B

Concentration dependence



C

Cleavage Site Sequence

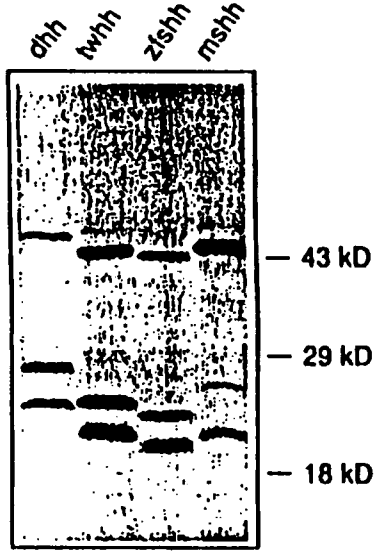
...SHVHG CFTPE...

253 262

D. melanogaster hh	ISSHVHGCFTPEST
D. hydei hh	SISHMHGCFTPEST
m- sonic hh	VAAKSGGCFPGSAT
r-sonic hh	VAAKSDGCFPGSAT
c-sonic hh	VAAKSGGCFPGSAL
z-sonic hh	VAAKSGGCFPGSGT
z-twhh'	VAAKSGGCFPAGAR
x-sonic hh	VAAKTGGCFPAGAQ
m-indian hh	VAAKTGGCFPGEAL
x-bhh	LGVRSGGCFPGTAM
x-hh4	LGVRSGGCFPGTAM
x-hh3	LGVRSGGCFPGTAM
m-desert hh	LAVRAGGCFPGNAT

D

In Vitro Translation (^{35}S-CYS)



E

Edman Degradation of *In vitro* translation Products

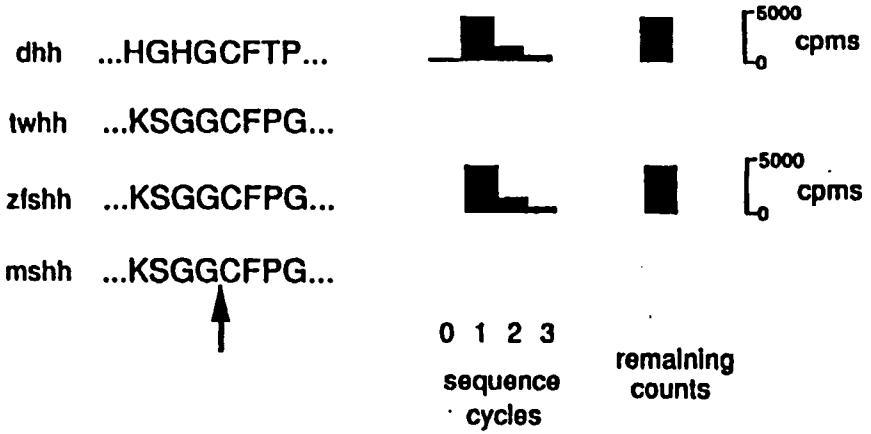


FIGURE 13

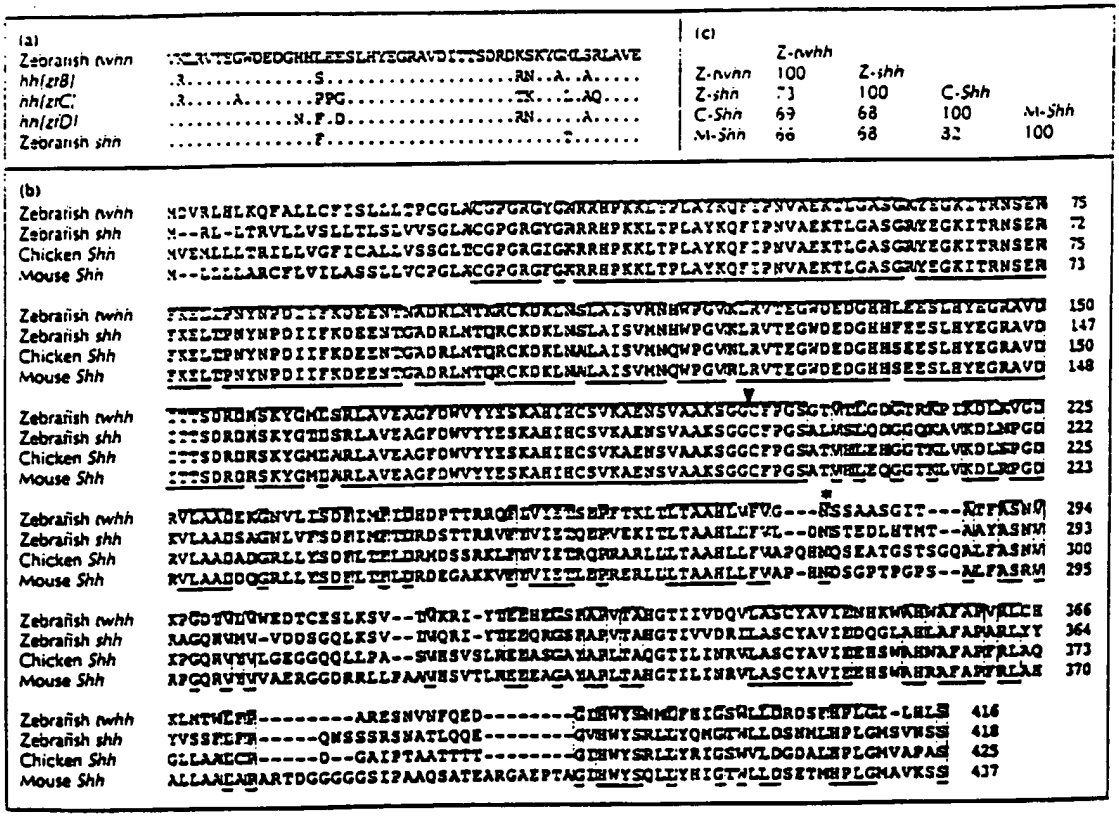
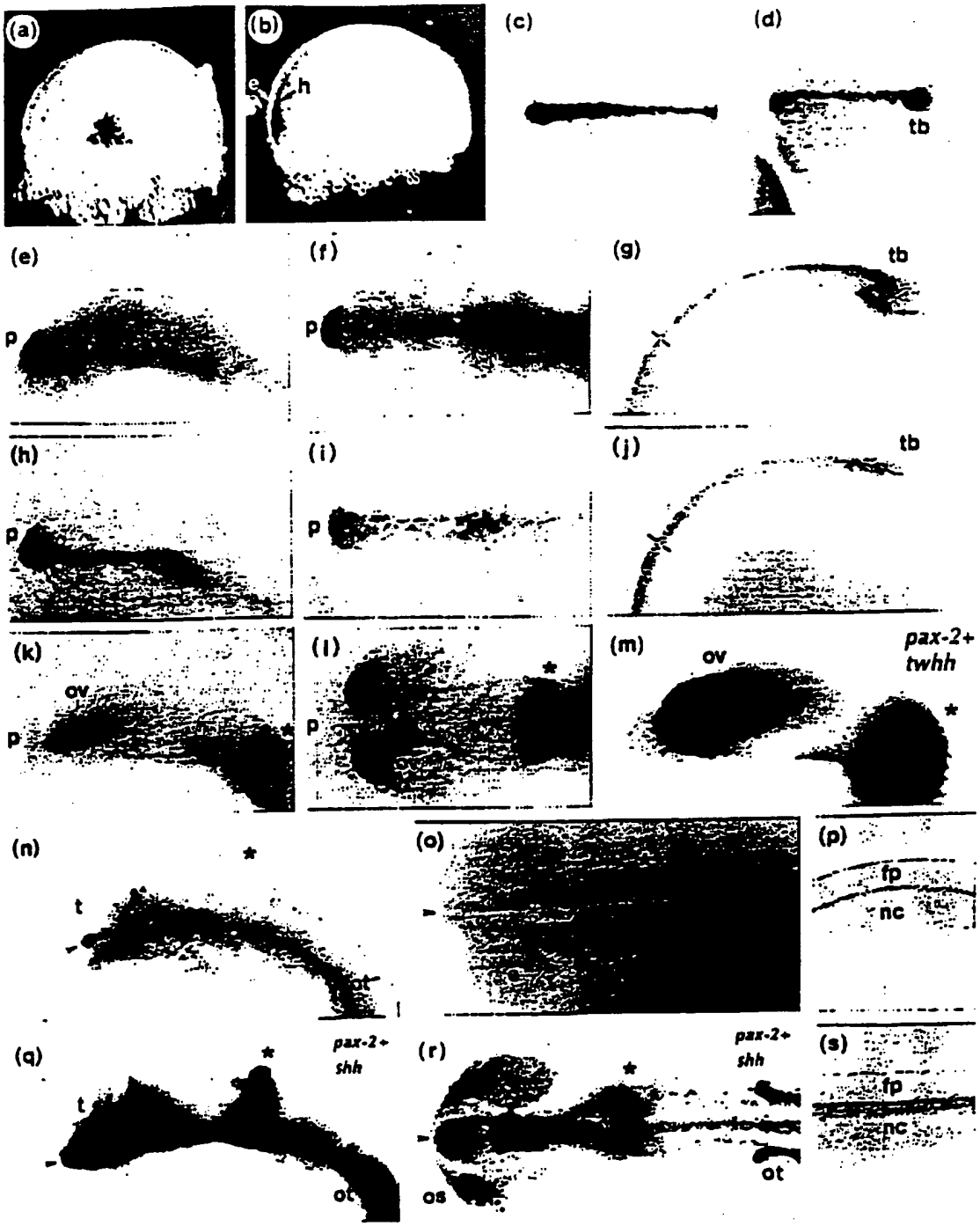

(a)
Zebrafish twhh
hh(zf8)
hh(zfC)
hh(zfD)
Zebrafish shh
(c)
Z-twhh
Z-twhh 100
Z-shh 73 100
C-Shh 69 68 100
M-Shh 66 68 32 100
Z-shh
C-Shh
M-Shh
(b)
Zebrafish twhh
Zebrafish shh
Chicken Shh
Mouse Shh
75
72
75
73
150
147
150
148
225
222
225
223
294
293
300
295
366
364
373
370
416
418
425
437

FIGURE 14


(a)
(b)
e
h
(c)
(d)
tb
(e)
p
(f)
p
(g)
tb
(h)
p
(i)
p
(j)
tb
(k)
ov
p
(l)
p
(m)
ov
pax-2+
twhh
(n)
t
ot
(o)
(p)
fp
nc
(q)
pax-2+
shh
t
ot
(r)
pax-2+
shh
os
ot
(s)
fp
nc

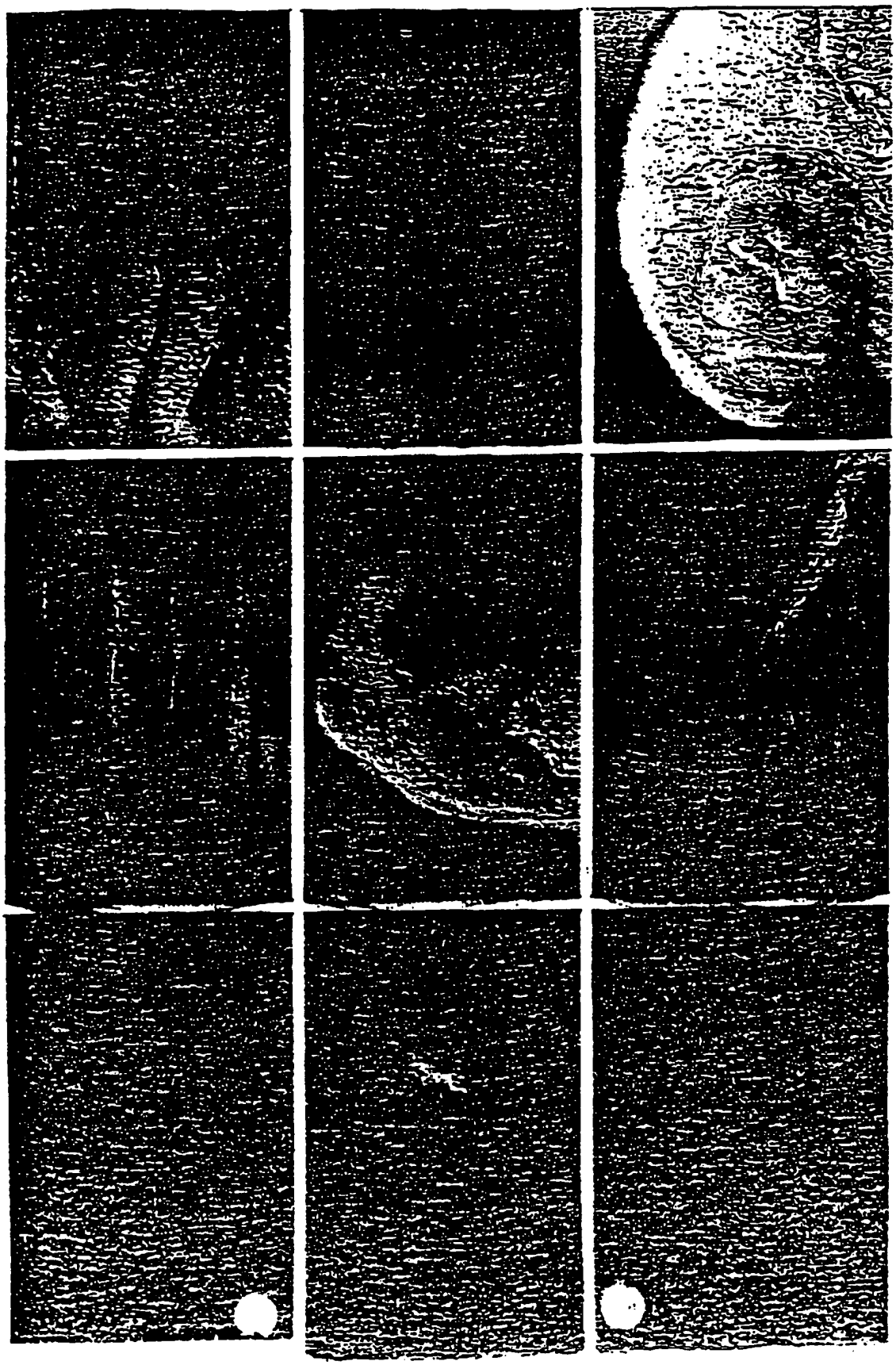


FIGURE 15

FIGURE 16

Table 1. Effects of ectopic expression of *shh*, *twhh*, *twhh-N*, *twhh-U*$_{HA}$ and *lac-Z* on zebrafish embryonic development.

Injected mRNA	*shh*	*twhh*	*twhh-N*	*twhh-U*$_{HA}$	*lac-Z*
12.5 h					
Ectopic *pax-2* in eye	89 % (35)	82 % (22)	92 % (26)	90 % (30)	0 % (31)
22 h					
Ectopic *pax-2* in eye	22 % (54)	62 % (50)	76 % (42)	21 % (39)	0 % (34)
Reduced *pax-6* in eye	20 %	68 %	54 %	1 %	0 %
Reduced *pax-6* in ventral forebrain	0 %	43 %	79 %	0 %	0 %
Reduced *pax-6* in hindbrain	0 % (35)	18 % (40)	61 % (28)	0 % (68)	0 % (14)
28 h					
Lens absent	16 %	86 %	100 %	9 %	0 %
Lens smaller	48 %	9 %	0 %	36 %	0 %
Reduced eye pigment	80 %	91 %	100 %	64 %	0 %
No midbrain–hindbrain constriction	48 % (25)	77 % (44)	100 % (16)	22 % (45)	3 % (37)

The percentages of affected embryos are shown; the numbers of embryos assayed are given in parentheses. Embryos were analyzed at the indicated time-points after injection of synthetic mRNA, as described in the text and Figs 3–5. These results represent a set of assays performed together for direct comparison of activities; repetitions yielded similar results. See text and Fig. 7 for description of transcription constructs.

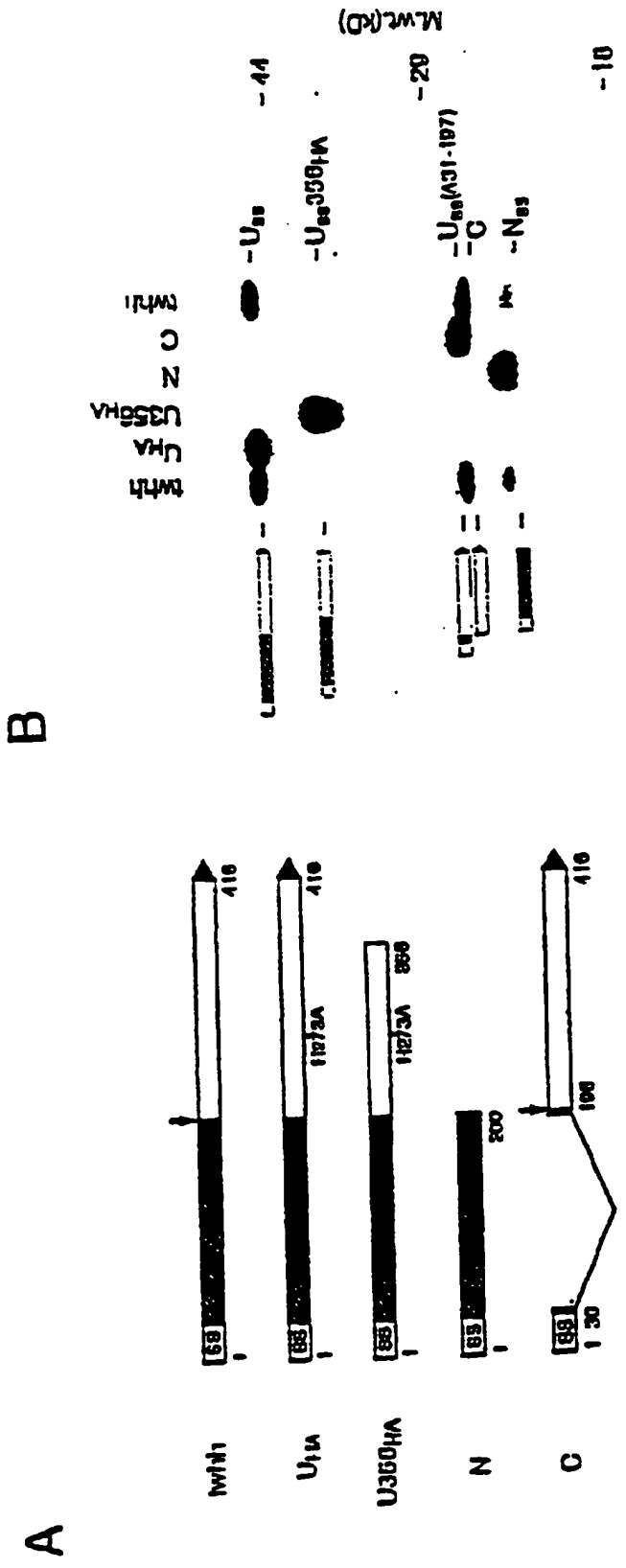


FIGURE 17

FIGURE 18

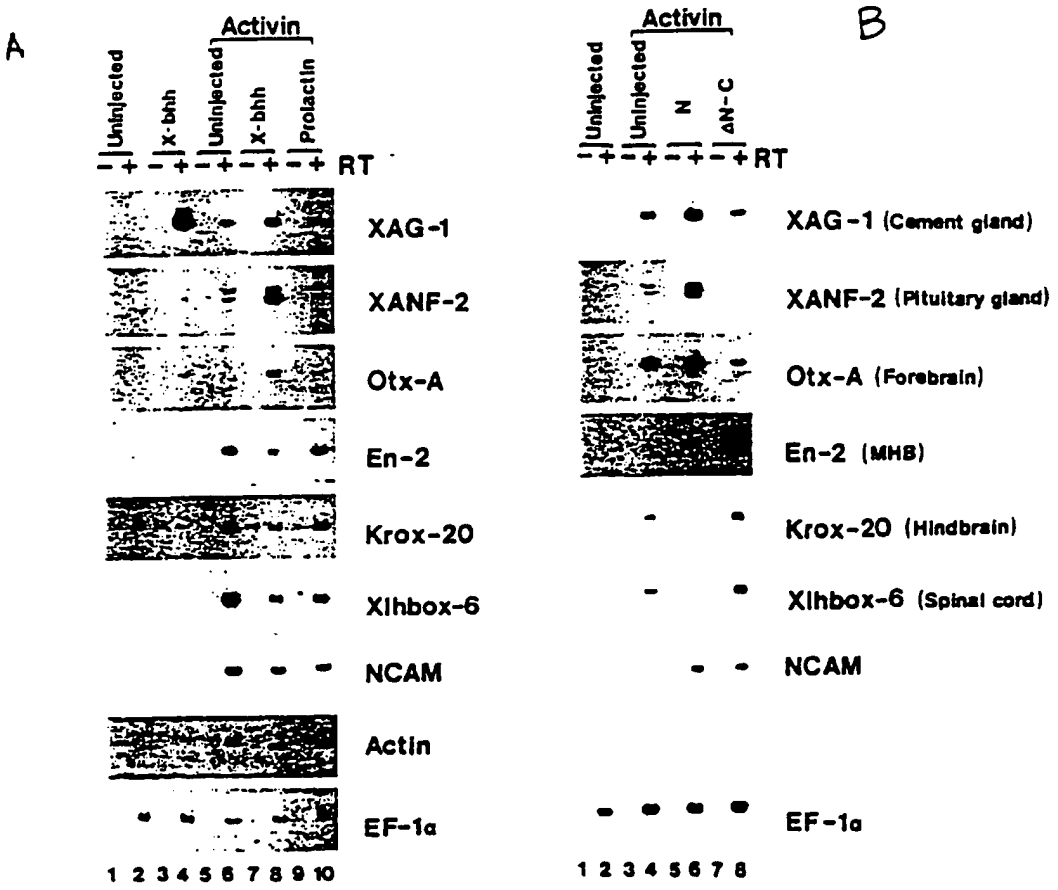

A
Activin
Uninjected
X-bhh
Uninjected
X-bhh
Prolactin
RT
XAG-1
XANF-2
Otx-A
En-2
Krox-20
Xlhbox-6
NCAM
Actin
EF-1α
1 2 3 4 5 6 7 8 9 10
B
Activin
Uninjected
Uninjected
N
ΔN-C
RT
XAG-1 (Cement gland)
XANF-2 (Pituitary gland)
Otx-A (Forebrain)
En-2 (MHB)
Krox-20 (Hindbrain)
Xlhbox-6 (Spinal cord)
NCAM
EF-1α
1 2 3 4 5 6 7 8

FIGURE 19

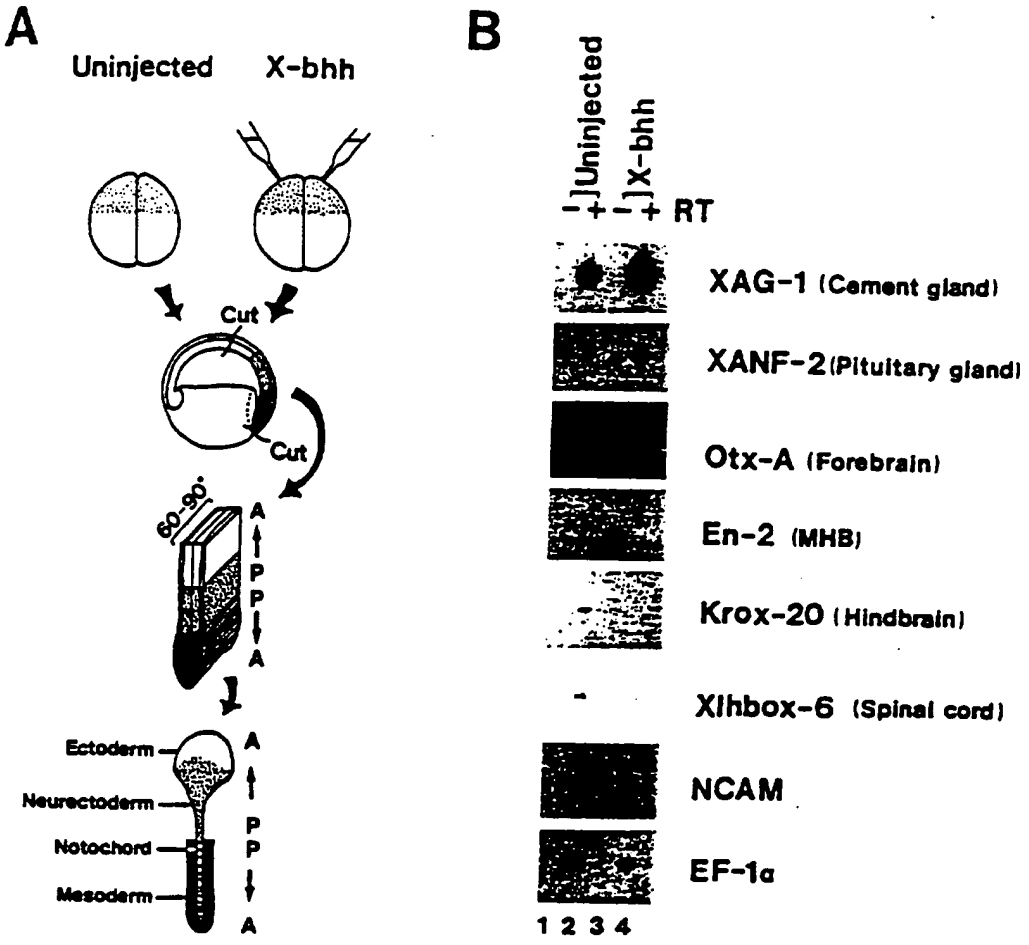
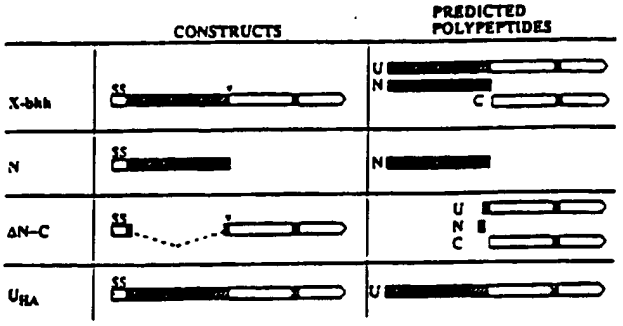

A
Uninjected
X-bhh
Cut
Cut
60-90°
A
P
P
A
Ectoderm
Neurectoderm
Notochord
Mesoderm
A
P
P
A
B
-
+
Uninjected
-
+
X-bhh
RT
XAG-1 (Cement gland)
XANF-2 (Pituitary gland)
Otx-A (Forebrain)
En-2 (MHB)
Krox-20 (Hindbrain)
Xlhbox-6 (Spinal cord)
NCAM
EF-1α
1 2 3 4

FIGURE 20


CONSTRUCTS
PREDICTED POLYPEPTIDES
X-bhh
N
ΔN-C
U
N
C

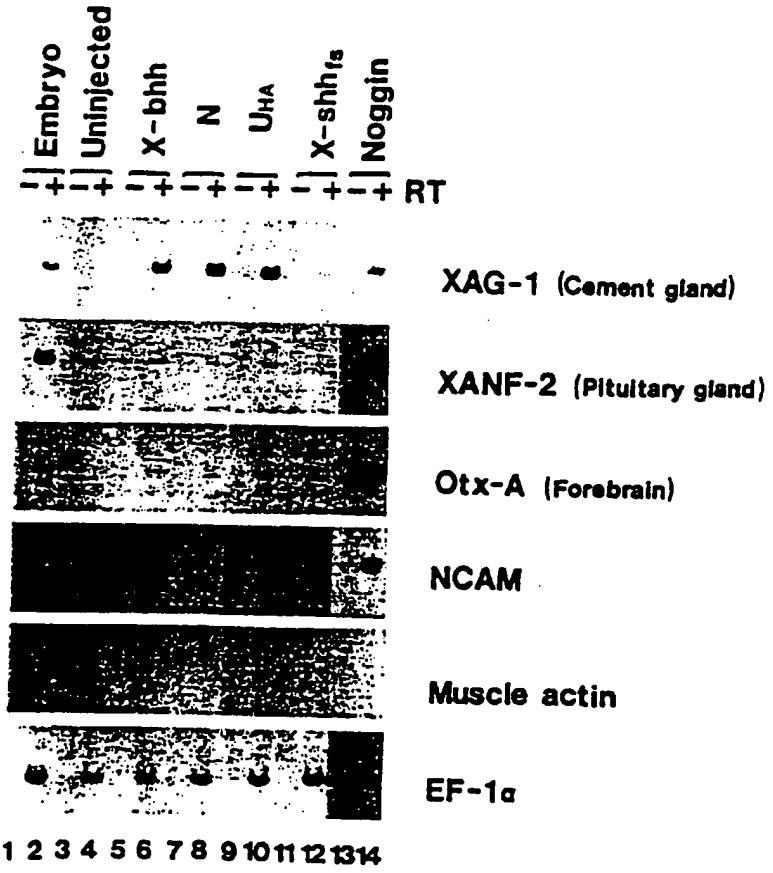
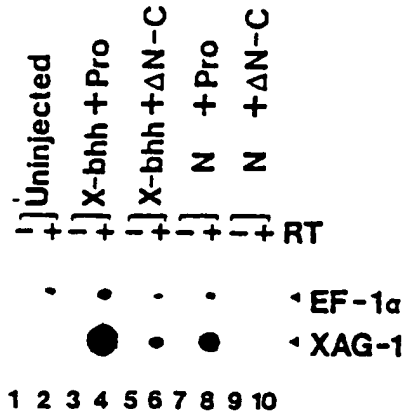

Embryo
Uninjected
X-bhh
N
UHA
X-shh1s
Noggin
RT
XAG-1 (Cement gland)
XANF-2 (Pituitary gland)
Otx-A (Forebrain)
NCAM
Muscle actin
EF-1α
1 2 3 4 5 6 7 8 9 10 11 12 13 14

FIGURE 21


Uninjected
X-bhh + Pro
X-bhh + ΔN-C
N + Pro
N + ΔN-C
RT
EF-1α
XAG-1
1 2 3 4 5 6 7 8 9 10

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : **A01N 61/00, A61K 38/16, 39/00, 39/395, C07H 21/02, 21/04, C12N 1/20, 15/00, C07K 16/00, G01N 33/53**

A3

(11) International Publication Number: **WO 96/17924**

(43) International Publication Date: 13 June 1996 (13.06.96)

(21) International Application Number: PCT/US95/15463

(22) International Filing Date: 4 December 1995 (04.12.95)

(30) Priority Data:
08/349,498 2 December 1994 (02.12.94) US

(71) Applicants: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US). UNIVERSITY OF WASHINGTON [US/US]; Suite 200, 1107 N.E. 45th Street, Seattle, WA 98105 (US).

(72) Inventors: BEACHY, Philip, A.; 5703 Chilham Road, Baltimore, MD 21209 (US). MOON, Randall, T.; 18531 57th Avenue N.W., Seattle, WA 98155 (US). PORTER, Jeffrey, A.; * (US).

(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: AL, AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, LS, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
26 September 1996 (26.09.96)

(54) Title: NOVEL HEDGEHOG-DERIVED POLYPEPTIDES

(57) Abstract

The present invention provides two novel polypeptides, referred to as the "N" and "C" fragments of hedgehog, or N-terminal and C-terminal fragments, respectively, which are derived after specific cleavage at a G↓CF site recognized by the autoproteolytic domain in the native protein. Also provided are methods of use of the N and C fragments.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15463

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/2; 530/358, 388.73, 389.1; 536/23.4, 23.5; 435/7.8, 252.3, 320.1; 424/130.1, 141.1, 158.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ------- Y	Lee et al. Autoproteolysis in hedgehog protein Biogenesis. Science. 02 December 1994, Vol. 266, pages 1528-1537, especially pages 1533-1535.	1-7, 9-15 ---------- 8, 16, 29-31
X ---- Y	Roelink et al. Floor Plate and Motor Neuron Induction by vhh-1, a Vertebrate Homolog of hedgehog Expressed by the Notochord. Cell. 25 February 1994, Vol. 76, pages 761-775, especially pages 767-769.	17-23 ---------- 24
X ---- Y	Basler et al. Control of cell pattern in the neural tube: regulation of cell differentiation by dorsalin-1, a novel TGFB family member. Cell. 21 May 1993, Vol 73, pages 687-702, especially pages 699-700.	25-27 -------- 28

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 11 JULY 1996	Date of mailing of the international search report 30 JUL 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer JOSEPH CURTIS Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/15463

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

A01N 61/00; A61K 38/16, 39/00, 39/395; C07H 21/02, 21/04; C12N 1/20, 15/00; C07K 16/00; G01N 33/53

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

514/2; 530/358, 388.73, 389.1; 536/23.4, 23.5; 435/7.8, 252.3, 320.1; 424/130.1, 141.1, 158.1

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, BIOSIS, CAPLUS, EMBASE, MEDLINE, WPIDS

search terms: hedgehog polypeptide, polynucleotide, DNA, N-terminal fragment, C-terminal fragment, antibody, method, neuronal proliferation, neuronal differentiation, expression vector.

Mediators of Hedgehog Signaling Pathways, Compositions and Uses Related Thereto

This application is based on U.S. Provisional Applications Nos. 60/154526, filed September 16, 1999, 60/159412, filed October 14, 1999, and 60/162899, filed November 1, 1999, the specifications of which are hereby incorporated by reference herein in their entirety.

Background of the Invention

Pattern formation is the activity by which embryonic cells form ordered spatial arrangements of differentiated tissues. The physical complexity of higher organisms arises during embryogenesis through the interplay of cell-intrinsic lineage and cell-extrinsic signaling. Inductive interactions are essential to embryonic patterning in vertebrate development from the earliest establishment of the body plan, to the patterning of the organ systems, to the generation of diverse cell types during tissue differentiation (Davidson, E., (1990) Development 108: 365-389; Gurdon, J. B., (1992) Cell 68: 185-199; Jessell, T. M. et al., (1992) Cell 68: 257-270). The effects of developmental cell interactions are varied. Typically, responding cells are diverted from one route of cell differentiation to another by inducing cells that differ from both the uninduced and induced states of the responding cells (inductions). Sometimes cells induce their neighbors to differentiate like themselves (homeogenetic induction); in other cases a cell inhibits its neighbors from differentiating like itself. Cell interactions in early development may be sequential, such that an initial induction between two cell types leads to a progressive amplification of diversity. Moreover, inductive interactions occur not only in embryos, but in adult cells as well, and can act to establish and maintain morphogenetic patterns as well as induce differentiation (J.B. Gurdon (1992) Cell 68:185-199).

Members of the *Hedgehog* family of signaling molecules mediate many important short- and long-range patterning processes during invertebrate and vertebrate development. In the fly, a single *hedgehog* gene regulates segmental and imaginal disc patterning. In contrast, in vertebrates, a *hedgehog* gene family is involved in the control of left-right asymmetry, polarity in the CNS, somites and limb, organogenesis, chondrogenesis and spermatogenesis.

The first *hedgehog* gene was identified by a genetic screen in the fruitfly *Drosophila melanogaster* (Nüsslein-Volhard, C. and Wieschaus, E. (1980) Nature 287, 795-801). This screen identified a number of mutations affecting embryonic and larval development. In 1992 and 1993, the molecular nature of the *Drosophila hedgehog* (*hh*) gene was reported (*C.F.*, Lee et al. (1992) Cell 71, 33-50), and since then, several *hedgehog* homologues have been isolated from various vertebrate species. While only one *hedgehog* gene has been found in *Drosophila* and other invertebrates, multiple *Hedgehog* genes are present in vertebrates.

The vertebrate family of *hedgehog* genes includes at least four members, e.g., paralogs of the single drosophila *hedgehog* gene. Exemplary hedgehog genes and proteins are described in PCT publications WO 95/18856 and WO 96/17924. Three of these members, herein referred to as Desert *hedgehog* (*Dhh*), Sonic *hedgehog* (*Shh*) and Indian *hedgehog* (*Ihh*), apparently exist in all vertebrates, including fish, birds, and mammals. A fourth member, herein referred to as tiggie-winkle *hedgehog* (*Thh*), appears specific to fish. Desert hedgehog (*Dhh*) is expressed principally in the testes, both in mouse embryonic development and in the adult rodent and human; Indian hedgehog (*Ihh*) is involved in bone development during embryogenesis and in bone formation in the adult; and, *Shh*, which as described above, is primarily involved in morphogenic and neuroinductive activities. Given the critical inductive roles of hedgehog polypeptides in the development and maintenance of vertebrate organs, the identification of hedghog interacting proteins is of paramount significance in both clinical and research contexts.

The various *Hedgehog* proteins consist of a signal peptide, a highly conserved N-terminal region, and a more divergent C-terminal domain. In addition to signal sequence

cleavage in the secretory pathway (Lee, J.J. et al. (1992) Cell 71:33-50; Tabata, T. *et al.* (1992) Genes Dev. 2635-2645; Chang, D.E. *et al.* (1994) Development 120:3339-3353), Hedgehog precursor proteins undergo an internal autoproteolytic cleavage which depends on conserved sequences in the C-terminal portion (Lee *et al.* (1994) Science 266:1528-1537; Porter *et al.* (1995) Nature 374:363-366). This autocleavage leads to a 19 kD N-terminal peptide and a C-terminal peptide of 26-28 kD (Lee *et al.* (1992) supra; Tabata *et al.* (1992) supra; Chang *et al.* (1994) supra; Lee *et al.* (1994) supra; Bumcrot, D.A., *et al.* (1995) Mol. Cell. Biol. 15:2294-2303; Porter *et al.* (1995) supra; Ekker, S.C. *et al.* (1995) Curr. Biol. 5:944-955; Lai, C.J. *et al.* (1995) Development 121:2349-2360). The N-terminal peptide stays tightly associated with the surface of cells in which it was synthesized, while the C-terminal peptide is freely diffusible both *in vitro* and *in vivo* (Porter et al. (1995) Nature 374:363; Lee *et al.* (1994) supra; Bumcrot *et al.* (1995) supra; Mart', E. *et al.* (1995) Development 121:2537-2547; Roelink, H. *et al.* (1995) Cell 81:445-455). Interestingly, cell surface retention of the N-terminal peptide is dependent on autocleavage, as a truncated form of HH encoded by an RNA which terminates precisely at the normal position of internal cleavage is diffusible *in vitro* (Porter *et al.* (1995) supra) and *in vivo* (Porter, J.A. *et al.* (1996) Cell 86, 21-34). Biochemical studies have shown that the autoproteolytic cleavage of the HH precursor protein proceeds through an internal thioester intermediate which subsequently is cleaved in a nucleophilic substitution. It is likely that the nucleophile is a small lipophilic molecule which becomes covalently bound to the C-terminal end of the N-peptide (Porter *et al.* (1996) supra), tethering it to the cell surface. The biological implications are profound. As a result of the tethering, a high local concentration of N-terminal Hedgehog peptide is generated on the surface of the Hedgehog producing cells. It is this N-terminal peptide which is both necessary and sufficient for short- and long-range Hedgehog signaling activities in *Drosophila* and vertebrates (Porter *et al.* (1995) supra; Ekker *et al.* (1995) supra; Lai *et al.* (1995) supra; Roelink, H. *et al.* (1995) Cell 81:445-455; Porter *et al.* (1996) supra; Fietz, M.J. *et al.* (1995) Curr. Biol. 5:643-651; Fan, C.-M. *et al.* (1995) Cell 81:457-465; Mart', E., *et al.* (1995) Nature 375:322-325; Lopez-Martinez *et al.* (1995) Curr. Biol

5:791-795; Ekker, S.C. *et al.* (1995) Development 121:2337-2347; Forbes, A.J. *et al.*(1996) Development 122:1125-1135).

HH has been implicated in short- and long-range patterning processes at various sites during *Drosophila* development. In the establishment of segment polarity in early embryos, it has short-range effects which appear to be directly mediated, while in the patterning of the imaginal discs, it induces long range effects via the induction of secondary signals.

In vertebrates, several *hedgehog* genes have been cloned in the past few years. Of these genes, *Shh* has received most of the experimental attention, as it is expressed in different organizing centers which are the sources of signals that pattern neighboring tissues. Recent evidence indicates that *Shh* is involved in these interactions.

The expression of *Shh* starts shortly after the onset of gastrulation in the presumptive midline mesoderm, the node in the mouse (Chang *et al.* (1994) supra; Echelard, Y. *et al.* (1993) Cell 75:1417-1430), the rat (Roelink, H. *et al.* (1994) Cell 76:761-775) and the chick (Riddle, R.D. *et al.* (1993) Cell 75:1401-1416), and the shield in the zebrafish (Ekker *et al.* (1995) supra; Krauss, S. *et al.*(1993) Cell 75:1431-1444). In chick embyros, the *Shh* expression pattern in the node develops a left-right asymmetry, which appears to be responsible for the left-right situs of the heart (Levin, M. *et al.* (1995) Cell 82:803-814).

In the CNS, *Shh* from the notochord and the floorplate appears to induce ventral cell fates. When ectopically expressed, *Shh* leads to a ventralization of large regions of the mid- and hindbrain in mouse (Echelard *et al.* (1993) supra; Goodrich, L.V. *et al.* (1996) Genes Dev. 10:301-312), *Xenopus* (Roelink, H. et al. (1994) *supra*; Ruiz i Altaba, A. *et al.* (1995) Mol. Cell. Neurosci. 6:106-121), and zebrafish (Ekker *et al.* (1995) supra; Krauss *et al.* (1993) supra; Hammerschmidt, M., *et al.* (1996) Genes Dev. 10:647-658). In explants of intermediate neuroectoderm at spinal cord levels, *Shh* protein induces floorplate and motor neuron development with distinct concentration thresholds, floor plate at high and motor neurons at lower concentrations (Roelink *et al.* (1995) supra; Mart' *et al.* (1995) supra; Tanabe, Y. *et al.* (1995) Curr. Biol. 5:651-658). Moreover,

antibody blocking suggests that *Shh* produced by the notochord is required for notochord-mediated induction of motor neuron fates (Mart' *et al.* (1995) supra). Thus, high concentration of *Shh* on the surface of *Shh*-producing midline cells appears to account for the contact-mediated induction of floorplate observed *in vitro* (Placzek, M. *et al.* (1993) Development 117:205-218), and the midline positioning of the floorplate immediately above the notochord *in vivo*. Lower concentrations of *Shh* released from the notochord and the floorplate presumably induce motor neurons at more distant ventrolateral regions in a process that has been shown to be contact-independent *in vitro* (Yamada, T. *et al.* (1993) Cell 73:673-686). In explants taken at midbrain and forebrain levels, *Shh* also induces the appropriate ventrolateral neuronal cell types, dopaminergic (Heynes, M. *et al.* (1995) Neuron 15:35-44; Wang, M.Z. *et al.* (1995) Nature Med. 1:1184-1188) and cholinergic (Ericson, J. *et al.* (1995) Cell 81:747-756) precursors, respectively, indicating that *Shh* is a common inducer of ventral specification over the entire length of the CNS. These observations raise a question as to how the differential response to *Shh* is regulated at particular anteroposterior positions.

Shh from the midline also patterns the paraxial regions of the vertebrate embryo, the somites in the trunk (Fan *et al.* (1995) supra) and the head mesenchyme rostral of the somites (Hammerschmidt *et al.* (1996) supra). In chick and mouse paraxial mesoderm explants, *Shh* promotes the expression of sclerotome specific markers like *Pax1* and *Twist*, at the expense of the dermamyotomal marker *Pax3*. Moreover, filter barrier experiments suggest that *Shh* mediates the induction of the sclerotome directly rather than by activation of a secondary signaling mechanism (Fan, C.-M. and Tessier-Lavigne, M. (1994) Cell 79, 1175-1186).

Shh also induces myotomal gene expression (Hammerschmidt et al. (1996) supra; Johnson, R.L. *et al.* (1994) Cell 79:1165-1173; Münsterberg, A.E. *et al.* (1995) Genes Dev. 9:2911-2922; Weinberg, E.S. *et al.* (1996) Development 122:271-280), although recent experiments indicate that members of the WNT family, vertebrate homologues of *Drosophila wingless*, are required in concert (Münsterberg *et al.* (1995) supra). Puzzlingly, myotomal induction in chicks requires higher *Shh* concentrations than the induction of sclerotomal markers (Münsterberg *et al.* (1995) supra), although the

sclerotome originates from somitic cells positioned much closer to the notochord. Similar results were obtained in the zebrafish, where high concentrations of Hedgehog induce myotomal and repress sclerotomal marker gene expression (Hammerschmidt *et al.* (1996) supra). In contrast to amniotes, however, these observations are consistent with the architecture of the fish embryo, as here, the myotome is the predominant and more axial component of the somites. Thus, modulation of *Shh* signaling and the acquisition of new signaling factors may have modified the somite structure during vertebrate evolution.

In the vertebrate limb buds, a subset of posterior mesenchymal cells, the "Zone of polarizing activity" (ZPA), regulates anteroposterior digit identity (reviewed in Honig, L.S. (1981) Nature 291:72-73). Ectopic expression of *Shh* or application of beads soaked in *Shh* peptide mimics the effect of anterior ZPA grafts, generating a mirror image duplication of digits (Chang *et al.* (1994) *supra*; Lopez-Martinez *et al.* (1995) *supra*; Riddle *et al.* (1993) *supra*) (Fig. 2g). Thus, digit identity appears to depend primarily on *Shh* concentration, although it is possible that other signals may relay this information over the substantial distances that appear to be required for AP patterning (100-150 μm). Similar to the interaction of HH and DPP in the *Drosophila* imaginal discs, *Shh* in the vertebrate limb bud activates the expression of *Bmp2* (Francis, P.H. *et al.* (1994) Development 120:209-218), a *dpp* homologue. However, unlike DPP in *Drosophila*, *Bmp2* fails to mimic the polarizing effect of *Shh* upon ectopic application in the chick limb bud (Francis *et al.* (1994) supra). In addition to anteroposterior patterning, *Shh* also appears to be involved in the regulation of the proximodistal outgrowth of the limbs by inducing the synthesis of the fibroblast growth factor FGF4 in the posterior apical ectodermal ridge (Laufer, E. *et al.* (1994) Cell 79:993-1003; Niswander, L. *et al.*(1994) Nature 371:609-612).

The close relationship between Hedgehog proteins and BMPs is likely to have been conserved at many, but probably not all sites of vertebrate *Hedgehog* expression. For example, in the chick hindgut, *Shh* has been shown to induce the expression of *Bmp4*, another vertebrate *dpp* homologue (Roberts, D.J. *et al.* (1995) Development 121:3163-3174). Furthermore, *Shh* and *Bmp2*, *4*, or *6* show a striking correlation in their expression in epithelial and mesenchymal cells of the stomach, the urogenital system, the lung, the

tooth buds and the hair follicles (Bitgood, M.J. and McMahon, A.P. (1995) Dev. Biol. 172:126-138). Further, *Ihh*, one of the two other mouse *Hedgehog* genes, is expressed adjacent to *Bmp* expressing cells in the gut and developing cartilage (Bitgood and McMahon (1995) *supra*).

Recent evidence suggests a model in which *Ihh* plays a crucial role in the regulation of chondrogenic development (Roberts *et al.* (1995) supra). During cartilage formation, chondrocytes proceed from a proliferating state via an intermediate, prehypertrophic state to differentiated hypertrophic chondrocytes. *Ihh* is expressed in the prehypertrophic chondrocytes and initiates a signaling cascade that leads to the blockage of chondrocyte differentiation. Its direct target is the perichondrium around the *Ihh* expression domain, which responds by the expression of *Gli* and *Patched* (*Ptc*), conserved transcriptional targets of Hedgehog signals (see below). Most likely, this leads to secondary signaling resulting in the synthesis of parathyroid hormone-related protein (PTHrP) in the periarticular perichondrium. PTHrP itself signals back to the prehypertrophic chondrocytes, blocking their further differentiation. At the same time, PTHrP represses expression of *Ihh*, thereby forming a negative feedback loop that modulates the rate of chondrocyte differentiation.

Patched was originally identified in Drosophila as a segment polarity gene, one of a group of developmental genes that affect cell differentiation within the individual segments that occur in a homologous series along the anterior-posterior axis of the embryo. See Hooper, J.E. et al. (1989) Cell 59:751; and Nakano, Y. et al. (1989) Nature 341:508. Patterns of expression of the vertebrate homologue of *patched* suggest its involvement in the development of neural tube, skeleton, limbs, craniofacial structure, and skin.

Genetic and functional studies demonstrate that *patched* is part of the hedgehog signaling cascade, an evolutionarily conserved pathway that regulates expression of a number of downstream genes. See Perrimon, N. (1995) Cell 80:517; and Perrimon, N. (1996) Cell 86:513. *Patched* participates in the constitutive transcriptional repression of the target genes; its effect is opposed by a secreted glycoprotein, encoded by hedgehog, or

a vertebrate homologue, which induces transcriptional activation. Genes under control of this pathway include members of the Wnt and TGF-beta families.

Patched proteins possess two large extracellular domains, twelve transmembrane segments, and several cytoplasmic segments. See Hooper, supra; Nakano, supra; Johnson, R.L. et al. (1996) Science 272:1668; and Hahn, H. et al. (1996) Cell 85:841. The biochemical role of *patched* in the hedgehog signaling pathway is unclear. Direct interaction with the hedgehog protein has, however, been reported (Chen, Y. et al. (1996) Cell 87:553), and *patched* may participate in a hedgehog receptor complex along with another transmembrane protein encoded by the *smoothened* gene. See Perrimon, supra; and Chen, supra.

The human homologue of *patched* was recently cloned and mapped to chromosome 9q22.3. See Johnson, supra; and Hahn, supra. This region has been implicated in basal cell nevus syndrome (BCNS), which is characterized by developmental abnormalities including rib and craniofacial alterations, abnormalities of the hands and feet, and spina bifida.

Sporadic tumors also demonstrated a loss of both functional alleles of *patched*. Of twelve tumors in which *patched* mutations were identified with a single strand conformational polymorphism screening assay, nine had chromosomal deletion of the second allele and the other three had inactivating mutations in both alleles (Gailani, supra). The alterations did not occur in the corresponding germline DNA.

Most of the identified mutations resulted in premature stop codons or frame shifts. Lench, N.J., et al., *Hum. Genet.* **1997** Oct; 100(5-6): 497-502. Several, however, were point mutations leading to amino acid substitutions in either extracellular or cytoplasmic domains. These sites of mutation may indicate functional importance for interaction with extracellular proteins or with cytoplasmic members of the downstream signaling pathway.

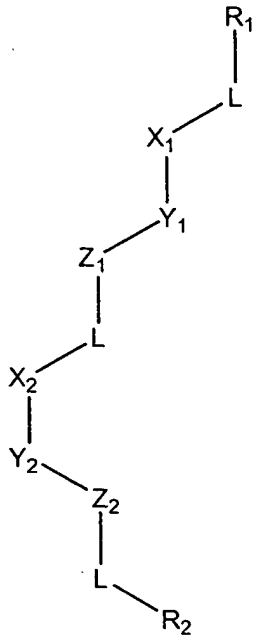
The involvement of *patched* in the inhibition of gene expression and the occurrence of frequent allelic deletions of *patched* in BCC support a tumor suppressor function for this gene. Its role in the regulation of gene families known to be involved in

cell signaling and intercellular communication provides a possible mechanism of tumor suppression.

Summary of the Invention

The present invention makes available methods and reagents for inhibiting aberrant growth states resulting from activation of the hedgehog signaling pathway, such as hedgehog gain-of-function, by contacting the cell with an agent, such as a small molecule, in a sufficient amount to reverse or control the aberrant growth state, e.g., to agonize a normal *ptc* pathway, antagonize a normal hedgehog pathway, or antagonize *smoothened* activity.

In one aspect, the invention pertains to a method for inhibiting an altered growth state of a cell having a *hedgehog* gain-of-function phenotype by contacting the cell with a *hedgehog* antagonist in a sufficient amount to inhibit the altered growth state, wherein the *hedgehog* antagonist is a organic molecule represented in the general formula (I):



Formula I

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X_1 and X_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-$N(R_8)$-, -$ON(R_8)$-, a heterocycle, or a direct bond between L and Y_1 or Y_2, respectively;

Y_1 and Y_2 are selected, independently, from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -P(=O)(OR_2)-, a heteroaromatic group, or a direct bond between X_1 and Z_1 or X_2 and Z_2, respectively;

Z_1 and Z_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y_1 or Y_2, respectively, and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), -$(CH_2)_n$heteroaryl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

p represents, independently for each occurrence, an integer from 0 to 10; and

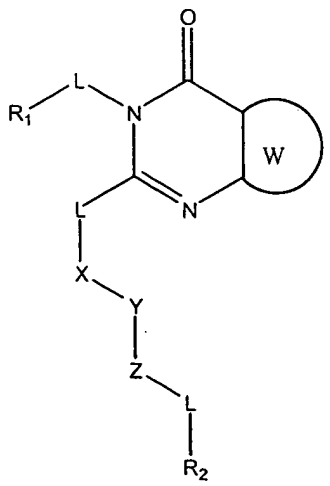
n, individually for each occurrence, represents an integer from 0 to 10.

In certain embodiments, R_1 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, at least one of X_1-Y_1-Z_1 and X_2-Y_2-Z_2 taken

together represents a urea or an amide. In certain embodiments, R_1 represents either a fused cycloalkyl-aryl or cycloalkyl-heteroaryl system.

In certain embodiments, the *hedgehog* antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 mM or less, 1 μM or less, 100 nM or less, 10 nM or less, or 1 nM or less. The cell may be contacted with the *hedgehog* antagonist *in vitro* or *in vivo*. In certain embodiments, the *hedgehog* antagonist is administered as part of a therapeutic or cosmetic application, such as regulation of neural tissues, bone and cartilage formation and repair, regulation of spermatogenesis, regulation of smooth muscle, regulation of lung, liver and other organs arising from the primative gut, regulation of hematopoietic function, or regulation of skin and hair growth.

In another aspect, the invention relates to a method for inhibiting an altered growth state of a cell having a *hedgehog* gain-of-function phenotype by contacting the cell with a *hedgehog* antagonist in a sufficient amount to inhibit the altered growth state, wherein the *hedgehog* antagonist is a organic molecule represented in the general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

X is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-(R_8)N-N(R_8)$-, $-ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y is selected from -C(=O)-, -C(=S)-, $-S(O_2)$-, -S(O)-, -C(=NCN)-, $-P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-R_8N-NR_8$-, $-ONR_8$-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10; and

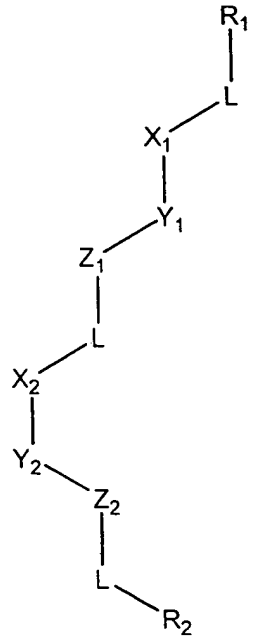
n, individually for each occurrence, represents an integer from 0 to 10.

In certain embodiments, R_1 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, X-Y-Z taken together represents a urea or an amide. In certain embodiments, W is a substituted or unsubstituted benzene ring. In certain embodiments, X represents a diazacarbocycle. In certain embodiments, R_2 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, R_8, independently for each occurrence, is selected from H and lower alkyl. In certain

embodiments, X is selected from -N(R_8)-, -O-, -S-, and a direct bond; Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-; and Z is selected from -N(R_8)-, -O-, -S-, and a direct bond, such that at least one of X and Z is present. In certain embodiments, at least one of X and Z is present. In certain embodiments, Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-.

In certain embodiments, the *hedgehog* antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 mM or less, 1 μM or less, 100 nM or less, 10 nM or less, or 1 nM or less. The cell may be contacted with the *hedgehog* antagonist *in vitro* or *in vivo*. In certain embodiments, the *hedgehog* antagonist is administered as part of a therapeutic or cosmetic application, such as regulation of neural tissues, bone and cartilage formation and repair, regulation of spermatogenesis, regulation of smooth muscle, regulation of lung, liver and other organs arising from the primative gut, regulation of hematopoietic function, or regulation of skin and hair growth.

In yet another embodiment, the invention provides a pharmaceutical preparation comprising a sterile pharmaceutical excipient and a compound represented by the general formula (I):



Formula I

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X_1 and X_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-$N(R_8)$-, -$ON(R_8)$-, a heterocycle, or a direct bond between L and Y_1 or Y_2, respectively;

Y_1 and Y_2 are selected, independently, from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -$P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X_1 and Z_1 or X_2 and Z_2, respectively;

Z_1 and Z_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y_1 or Y_2, respectively, and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), -$(CH_2)_n$heteroaryl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

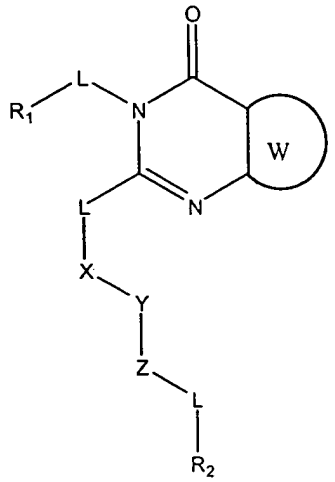
p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

In certain embodiments, R_1 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, at least one of X_1-Y_1-Z_1 and X_2-Y_2-Z_2 taken

together represents a urea or an amide. In certain embodiments, R_1 represents either a fused cycloalkyl-aryl or cycloalkyl-heteroaryl system.

In another aspect, the invention relates to a pharmaceutical preparation comprising a sterile pharmaceutical excipient and a compound represented by the general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

X is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-(R_8)N$-$N(R_8)$-, $-ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y is selected from -C(=O)-, -C(=S)-, $-S(O_2)$-, -S(O)-, -C(=NCN)-, $-P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

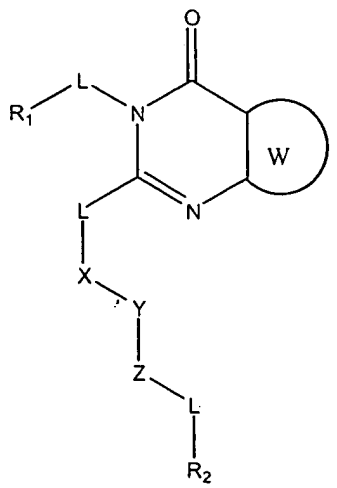
W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

In certain embodiments, R_1 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, X-Y-Z taken together represents a urea or an amide. In certain embodiments, W is a substituted or unsubstituted benzene ring. In certain embodiments, X represents a diazacarbocycle. In certain embodiments, R_2 represents a substituted or unsubstituted aryl or heteroaryl group. In certain embodiments, R_8, independently for each occurrence, is selected from H and lower alkyl. In certain embodiments, X is selected from -N(R_8)-, -O-, -S-, and a direct bond; Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-; and Z is selected from -N(R_8)-, -O-, -S-, and a direct bond, such that at least one of X and Z is present. In certain embodiments, at least one of X and Z is present. In certain embodiments, Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-.

In still another aspect, the invention relates to a compound having the general structure of Formula II:



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X is -NH-;

Y is selected from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -$P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted), or two R_8 taken together may form a 4- to 8-membered ring, e.g., with X_1 and Z_1 or X_2 and Z_2, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

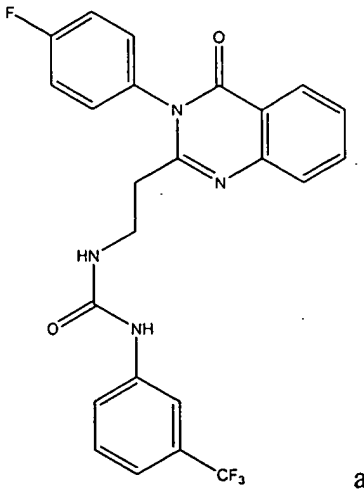
p represents, independently for each occurrence, an integer from 0 to 10, preferably from 0 to 3; and

n, individually for each occurence, represents an integer from 0 to 10, preferably from 0 to 5.

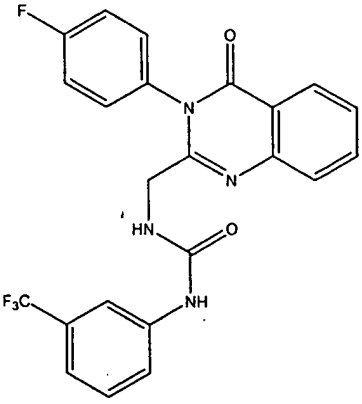
In certain embodiments, L adjacent to X represents -(unbranched lower alkyl)-. In certain embodiments, R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, aryl, hydroxyl, (unbranched alkyl-O-), silyloxy, amino, nitro, thiol, amino, imino, amido, phosphoryl, phosphonate, phosphine, carbonyl, carboxyl, carboxamide, anhydride, silyl, thioether, alkylsulfonyl, arylsulfonyl, sulfoxide, selenoether, ketone, aldehyde, ester, or $-(CH_2)_m-R_8$.

In certain embodiments, R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, aryl, nitro, thiol, imino, amido, carbonyl, carboxyl, anhydride, thioether, alkylsulfonyl, arylsulfonyl, ketone, aldehyde, and ester. In certain embodiments, R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, nitro, amido, carboxyl, anhydride, alkylsulfonyl, ketone, aldehyde, and ester. In certain embodiments, R_2 is a substituted or unsubstituted aryl or heteroaryl ring. In certain embodiments, X-Y-Z- represents an amide or urea linkage. In certain embodiments, R_8 represents H for all occurrences. In certain embodiments, R_2 is a substituted or unsubstituted aryl or heteroaryl ring. In certain embodiments, L is absent adjacent to R_1.

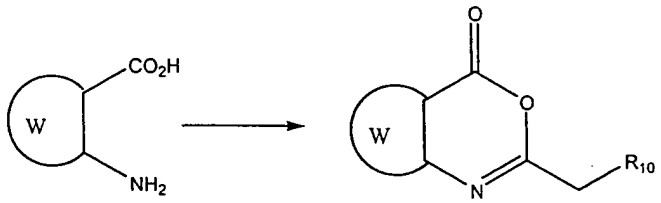
In certain embodiments, the compound has a structure selected from the structures depicted in Figures 32j, k, and l,



and



In yet another aspect, the invention relates to a method for preparing a bicyclic compound, comprising heating a reaction mixture comprising a compound having a structure of Formula X with a carboxylic acid anhydride according to the scheme:



X XI

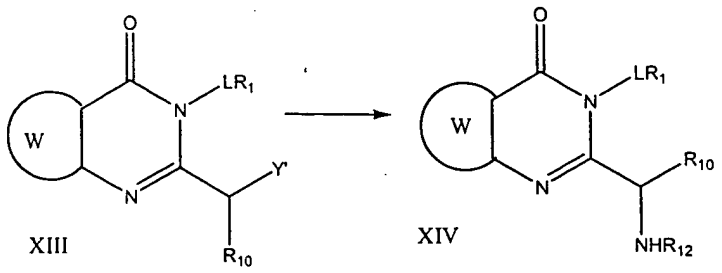
wherein W represents an aryl or heteroaryl ring, and

R_{10} represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl, and

wherein the anhydride has a structure of $(R_{10}CH_2C{=}O)_2O$, wherein R' is CH_2R_{10}.

In certain embodiments, the reaction mixture consists essentially of the compound of Formula X and the carboxylic acid anhydride at the start of the reaction.

In still another aspect, the invention provides a method for preparing an amine, comprising contacting a compound having a structure of Formula XIII with an amine having a structure of H_2NR_{12} according to the scheme:



wherein W represents an aryl or heteroaryl ring;

R_1 represents H, lower alkyl, $-(CH_2)_n$aryl, or $-(CH_2)_n$heteroaryl;

L is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

Y' represents a halogen;

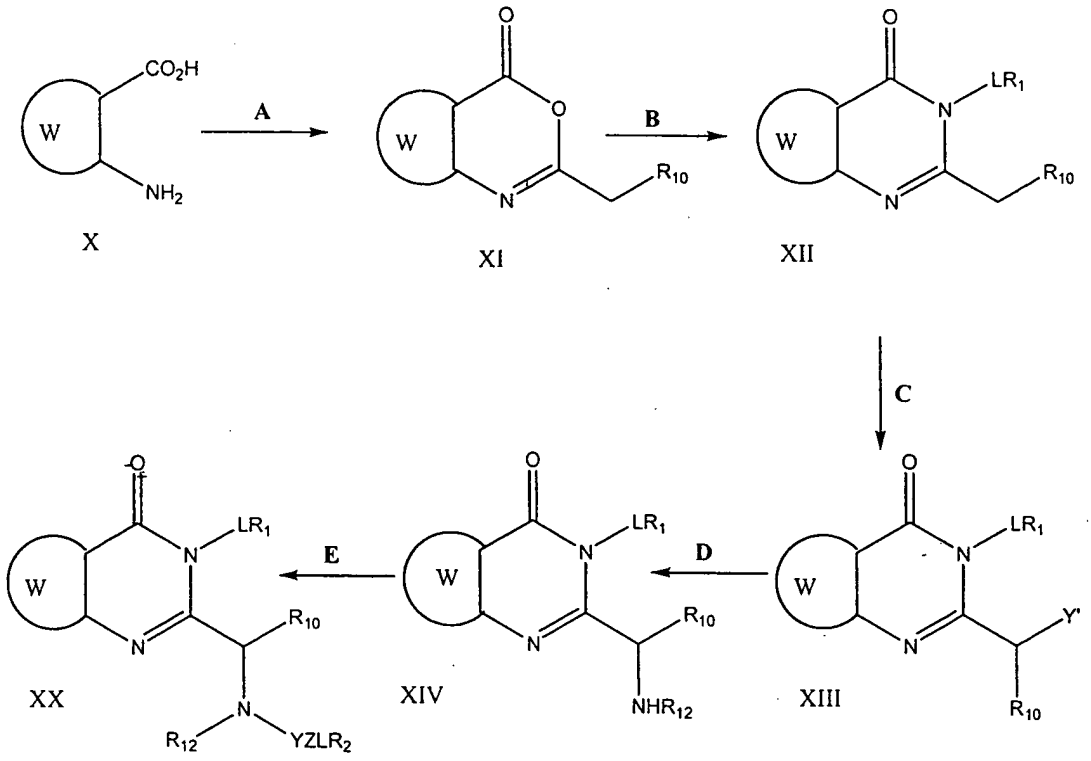
R_{10} represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl; and

R_{12} represents a lower alkyl group or a silyl group,

wherein the amine and the compound having a structure of Formula XIII are combined with a polar solvent comprising less than about 50% water.

In certain embodiments, the polar solvent comprises an alcohol. In certain embodiments, the alcohol is selected from methanol, ethanol, propanol, isopropanol, butanol, isobutanol, t-butanol, sec-butanol, ethylene glycol, and 1,3-propanediol.

In another aspect, the invention provides a method for preparing a compound, comprising performing steps according to the scheme:



wherein step (A) comprises reacting a compound having a structure of Formula X, wherein W represents a substituted or unsubstituted aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid group in adjacent (*ortho*) positions, with an acylating agent having the formula $R_{10}CH_2C(=O)X'$, wherein R_{10}, independently for each occurrence, represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl, and X' represents a halogen or $-OC(=O)CH_2R_{10}$, under conditions that produce a compound having a structure of Formula XI;

step (B) comprises reacting a compound having a structure of Formula XI with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XII;

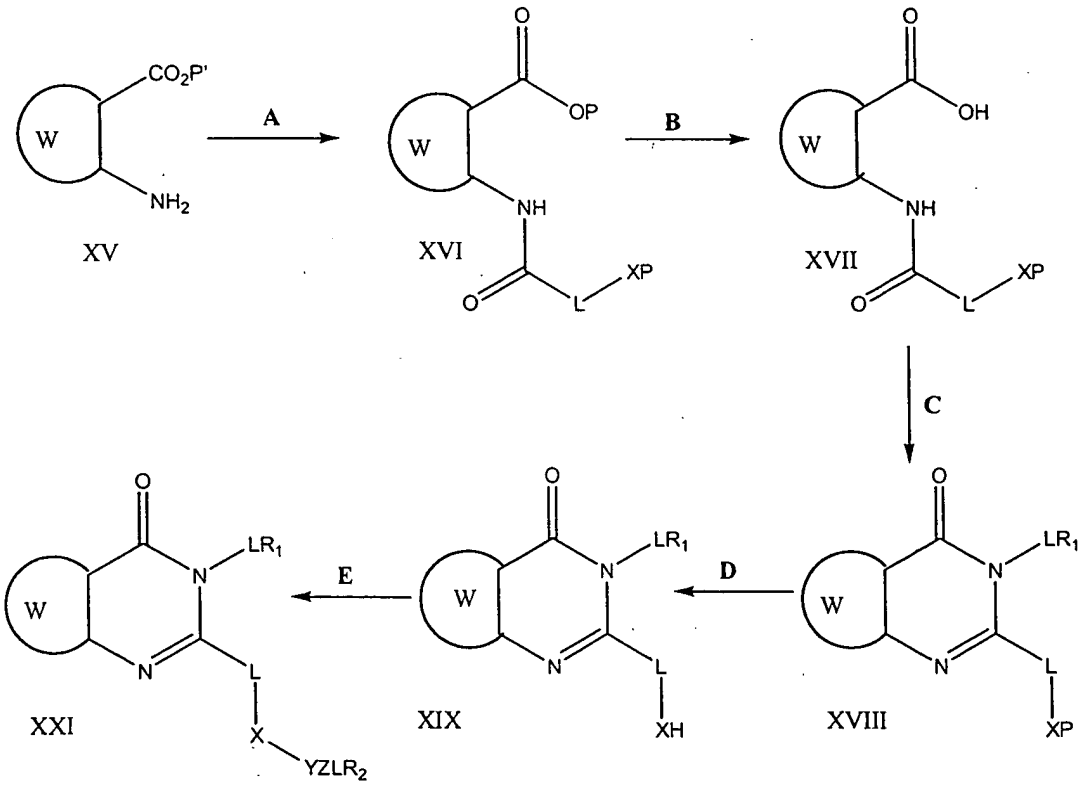
step (C) comprises reacting a compound having a structure of Formula XII with a halogenating agent, such as chlorine, bromine, iodine, N-bromosuccinimide, N-chlorosuccinimide, N-iodosuccinimide, ClBr, IBr, ClI, or a reagent that generates a

halogen radical (such as $Cl\cdot$, $Br\cdot$, or $I\cdot$) under conditions that result in a compound having a structure of Formula XIII, wherein Y' represents a halogen such as Cl, Br, or I;

step (D) comprises reacting a compound having a structure of Formula XIII with an amine having the formula H_2NR_{12}, wherein R_{12} represents a lower alkyl group or a silyl group, such as a trialkylsilyl, triarylsilyl, dialkylarylsilyl, or diarylalkylsilyl group, under conditions that result in a compound having a structure of Formula XIV; and

step (E) comprises reacting a compound having a structure of Formula XIV with a terminating group having a structure of R_2V' to produce a compound having a structure of Formula XX, wherein R_2 is as defined above, and V' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, ZC(=W)WC(=W)ZR_2, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an activated acylating moiety prepared *in situ*.

In still another aspect, the invention provides a method for preparing a compound, comprising performing steps according to the scheme:



wherein step (A) comprises reacting a compound having a structure of Formula XV, wherein P' represents H or a protecting group, W represents a substituted or unsubstituted

aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid or ester group in adjacent (*ortho*) positions, with an acylating agent having the formula PXLC(=O)X', wherein X and L are as defined above, P represents a protecting group, and X' represents a halogen, –OC(=O)LXP, or a functional group generated by reacting a carboxyl group with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, carbonyldiimidazole, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an increased susceptibility, relative to the carboxylic acid, towards coupling with an amine, under conditions that produce a compound having a structure of Formula XVI;

step (B) comprises deprotecting the ester of a compound having a structure of Formula XVI to produce a carboxylic acid having a structure of Formula XVII, if necessary;

step (C) comprises reacting a compound having a structure of Formula XVII with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XVIII;

step (D) comprises removing the protecting group P from a compound having a structure of Formula XVIII to generate a compound having a structure of Formula XIX;

step (E) comprises reacting a compound having a structure of Formula XIX with a terminating group having a structure of R_2Y' to produce a compound having a structure of Formula XXI, wherein R_2 is as defined above, and Y' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, $ZC(=W)WC(=W)ZR_2$, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an active acylating moiety prepared *in situ*.

Brief Description of the Drawings

Figures 1-31 depict reactions useful for synthesizing compounds according to the present invention.

Figures 32a-l illustrate representative compounds according to the present invention.

Figure 33 presents biological testing results of an exemplary compound of the invention.

Figure 34 presents the effects of an exemplary compound of the invention on lung tissue.

Detailed Description of the Invention

I. Overview

The present invention relates to the discovery that signal transduction pathways regulated by *hedgehog, patched* (*ptc*), *gli* and/or *smoothened* can be inhibited, at least in part, by small molecules. While not wishing to bound by any particular theory, the activation of a receptor may be the mechanism by which these agents act. For example, the ability of these agents to inhibit proliferation of *patched* loss-of-function (ptc^{lof}) cells may be due to the ability of such molecules to interact with *hedgehog*, *patched*, or *smoothened*, or at least to interfere with the ability of those proteins to activate a *hedgehog*, *ptc*, and/or *smoothened*-mediated signal transduction pathway.

It is, therefore, specifically contemplated that these small molecules which intefere with aspects of *hedgehog*, *ptc*, or *smoothened* signal transduction activity will likewise be capable of inhibiting proliferation (or other biological consequences) in cells having a *hedgehog* gain-of-function phenotype. In preferred embodiments, the subject inhibitors are organic molecules having a molecular weight less than 2500 amu, more preferably less than 1500 amu, and even more preferably less than 750 amu, and are capable of inhibiting at least some of the biological activities of hedgehog proteins, preferably specifically in target cells.

Thus, the methods of the present invention include the use of small molecules which agonize *ptc* inhibition of *hedgehog* signalling in the regulation of repair and/or functional performance of a wide range of cells, tissues and organs having the phenotype of *hedgehog* gain-of-function. For instance, the subject method has therapeutic and cosmetic applications ranging from regulation of neural tissues, bone and cartilage formation and repair, regulation of spermatogenesis, regulation of smooth muscle, regulation of lung, liver and tissue of other organs arising from the primative gut,

regulation of hematopoietic function, regulation of skin and hair growth, etc. Moreover, the subject methods can be performed on cells which are provided in culture (*in vitro*), or on cells in a whole animal (*in vivo*). See, for example, PCT publications WO 95/18856 and WO 96/17924 (the specifications of which are expressly incorporated by reference herein).

In a preferred embodiment, the subject method can be to treat epithelial cells having a phenotype of *hedgehog* gain-of-function.

In another preferred embodiment, the subject method can be used as part of a treatment regimen for malignant medulloblastoma and other primary CNS malignant neuroectodermal tumors.

In another aspect, the present invention provides pharmaceutical preparations comprising, as an active ingredient, a *hedgehog* antagonist or *ptc* agonist such as described herein, formulated in an amount sufficient to inhibit, *in vivo*, proliferation or other biological consequences of *hedgehog* gain-of-function.

The subject treatments using *hedgehog* antagonists can be effective for both human and animal subjects. Animal subjects to which the invention is applicable extend to both domestic animals and livestock, raised either as pets or for commercial purposes. Examples are dogs, cats, cattle, horses, sheep, hogs, and goats.

II. Definitions

For convience, certain terms employed in the specification, examples, and appended claims are collected here.

The phrase "aberrant modification or mutation" of a gene refers to such genetic lesions as, for example, deletions, substitution or addition of nucleotides to a gene, as well as gross chromosomal rearrangements of the gene and/or abnormal methylation of the gene. Likewise, mis-expression of a gene refers to aberrant levels of transcription of the gene relative to those levels in a normal cell under similar conditions, as well as non-wild-type splicing of mRNA transcribed from the gene.

"Basal cell carcinomas" exist in a variety of clinical and histological forms such as nodular-ulcerative, superficial, pigmented, morphealike, fibroepithelioma and nevoid syndrome. Basal cell carcinomas are the most common cutaneous neoplasms found in humans. The majority of new cases of nonmelanoma skin cancers fall into this category.

"Burn wounds" refer to cases where large surface areas of skin have been removed or lost from an individual due to heat and/or chemical agents.

The term "carcinoma" refers to a malignant new growth made up of epithelial cells tending to infiltrate surrounding tissues and to give rise to metastases. Exemplary carcinomas include: "basal cell carcinoma", which is an epithelial tumor of the skin that, while seldom metastasizing, has potentialities for local invasion and destruction; "squamous cell carcinoma", which refers to carcinomas arising from squamous epithelium and having cuboid cells; "carcinosarcoma", which include malignant tumors composed of carcinomatous and sarcomatous tissues; "adenocystic carcinoma", carcinoma marked by cylinders or bands of hyaline or mucinous stroma separated or surrounded by nests or cords of small epithelial cells, occurring in the mammary and salivary glands, and mucous glands of the respiratory tract; "epidermoid carcinoma", which refers to cancerous cells which tend to differentiate in the same way as those of the epidermis; i.e., they tend to form prickle cells and undergo cornification; "nasopharyngeal carcinoma", which refers to a malignant tumor arising in the epithelial lining of the space behind the nose; and "renal cell carcinoma", which pertains to carcinoma of the renal parenchyma composed of tubular cells in varying arrangements. Other carcinomatous epithelial growths are "papillomas", which refers to benign tumors derived from epithelium and having a papillomavirus as a causative agent; and "epidermoidomas", which refers to a cerebral or meningeal tumor formed by inclusion of ectodermal elements at the time of closure of the neural groove.

The "corium" or "dermis" refers to the layer of the skin deep to the epidermis, consisting of a dense bed of vascular connective tissue, and containing the nerves and terminal organs of sensation. The hair roots, and sebaceous and sweat glands are structures of the epidermis which are deeply embedded in the dermis.

"Dental tissue" refers to tissue in the mouth which is similar to epithelial tissue, for example gum tissue. The method of the present invention is useful for treating periodontal disease.

"Dermal skin ulcers" refer to lesions on the skin caused by superficial loss of tissue, usually with inflammation. Dermal skin ulcers which can be treated by the method of the present invention include decubitus ulcers, diabetic ulcers, venous stasis ulcers and arterial ulcers. Decubitus wounds refer to chronic ulcers that result from pressure applied to areas of the skin for extended periods of time. Wounds of this type are often called bedsores or pressure sores. Venous stasis ulcers result from the stagnation of blood or other fluids from defective veins. Arterial ulcers refer to necrotic skin in the area around arteries having poor blood flow.

The term "ED_{50}" means the dose of a drug which produces 50% of its maximum response or effect.

An "effective amount" of, e.g., a *hedgehog* antagonist, with respect to the subject method of treatment, refers to an amount of the antagonist in a preparation which, when applied as part of a desired dosage regimen brings about, e.g., a change in the rate of cell proliferation and/or the state of differentiation of a cell and/or rate of survival of a cell according to clinically acceptable standards for the disorder to be treated or the cosmetic purpose.

The terms "epithelia", "epithelial" and "epithelium" refer to the cellular covering of internal and external body surfaces (cutaneous, mucous and serous), including the glands and other structures derived therefrom, e.g., corneal, esophegeal, epidermal, and hair follicle epithelial cells. Other exemplary epithlelial tissue includes: olfactory epithelium, which is the pseudostratified epithelium lining the olfactory region of the nasal cavity, and containing the receptors for the sense of smell; glandular epithelium, which refers to epithelium composed of secreting cells; squamous epithelium, which refers to epithelium composed of flattened plate-like cells. The term epithelium can also refer to transitional epithelium, like that which is characteristically found lining hollow organs that are subject to great mechanical change due to contraction and distention, e.g.,

tissue which represents a transition between stratified squamous and columnar epithelium.

The term "epithelialization" refers to healing by the growth of epithelial tissue over a denuded surface.

The term "epidermal gland" refers to an aggregation of cells associated with the epidermis and specialized to secrete or excrete materials not related to their ordinary metabolic needs. For example, "sebaceous glands" are holocrine glands in the corium that secrete an oily substance and sebum. The term "sweat glands" refers to glands that secrete sweat, situated in the corium or subcutaneous tissue, opening by a duct on the body surface.

The term "epidermis" refers to the outermost and nonvascular layer of the skin, derived from the embryonic ectoderm, varying in thickness from 0.07-1.4 mm. On the palmar and plantar surfaces it comprises, from within outward, five layers: basal layer composed of columnar cells arranged perpendicularly; prickle-cell or spinous layer composed of flattened polyhedral cells with short processes or spines; granular layer composed of flattened granular cells; clear layer composed of several layers of clear, transparent cells in which the nuclei are indistinct or absent; and horny layer composed of flattened, cornified non-nucleated cells. In the epidermis of the general body surface, the clear layer is usually absent.

"Excisional wounds" include tears, abrasions, cuts, punctures or lacerations in the epithelial layer of the skin and may extend into the dermal layer and even into subcutaneous fat and beyond. Excisional wounds can result from surgical procedures or from accidental penetration of the skin.

The "growth state" of a cell refers to the rate of proliferation of the cell and/or the state of differentiation of the cell. An "altered growth state" is a growth state characterized by an abnormal rate of proliferation, e.g., a cell exhibiting an increased or decreased rate of proliferation relative to a normal cell.

The term "hair" refers to a threadlike structure, especially the specialized epidermal structure composed of keratin and developing from a papilla sunk in the

corium, produced only by mammals and characteristic of that group of animals. Also, "hair" may refer to the aggregate of such hairs. A "hair follicle" refers to one of the tubular-invaginations of the epidermis enclosing the hairs, and from which the hairs grow. "Hair follicle epithelial cells" refers to epithelial cells which surround the dermal papilla in the hair follicle, e.g., stem cells, outer root sheath cells, matrix cells, and inner root sheath cells. Such cells may be normal non-malignant cells, or transformed/immortalized cells.

The term "*hedgehog* antagonist" refers to an agent which potentiates or recapitulates the bioactivity of *patched*, such as to repress transcription of target genes. Preferred *hedgehog* antagonists can be used to overcome a *ptc* loss-of-function and/or a *smoothened* gain-of-function, the latter also being refered to as *smoothened* antagonists. The term '*hedgehog* antagonist' as used herein refers not only to any agent that may act by directly inhibiting the normal function of the *hedgehog* protein, but also to any agent that inhibits the hedgehog signalling pathway, and thus recapitulates the function of *ptc*.

The term "*hedgehog* gain-of-function" refers to an aberrant modification or mutation of a *ptc* gene, *hedgehog* gene, or *smoothened* gene, or a decrease (or loss) in the level of expression of such a gene, which results in a phenotype which resembles contacting a cell with a hedgehog protein, e.g., aberrant activation of a hedgehog pathway. The gain-of-function may include a loss of the ability of the *ptc* gene product to regulate the level of expression of Ci genes, e.g., *Gli1*, *Gli2*, and *Gli3*. The term '*hedgehog* gain-of-function' is also used herein to refer to any similar cellular phenotype (e.g., exhibiting excess proliferation) which occurs due to an alteration anywhere in the hedgehog signal transduction pathway, including, but not limited to, a modification or mutation of *hedgehog* itself. For example, a tumor cell with an abnormally high proliferation rate due to activation of the hedgehog signalling pathway would have a '*hedgehog* gain-of-function' phenotype, even if *hedgehog* is not mutated in that cell.

As used herein, "immortalized cells" refers to cells which have been altered via chemical and/or recombinant means such that the cells have the ability to grow through an indefinite number of divisions in culture.

"Internal epithelial tissue" refers to tissue inside the body which has characteristics similar to the epidermal layer in the skin. Examples include the lining of the intestine. The method of the present invention is useful for promoting the healing of certain internal wounds, for example wounds resulting from surgery.

The term "keratosis" refers to proliferative skin disorder characterized by hyperplasia of the horny layer of the epidermis. Exemplary keratotic disorders include keratosis follicularis, keratosis palmaris et plantaris, keratosis pharyngea, keratosis pilaris, and actinic keratosis.

The term "LD_{50}" means the dose of a drug which is lethal in 50% of test subjects.

The term "nail" refers to the horny cutaneous plate on the dorsal surface of the distal end of a finger or toe.

The term "*patched* loss-of-function" refers to an aberrant modification or mutation of a *ptc* gene, or a decreased level of expression of the gene, which results in a phenotype which resembles contacting a cell with a hedgehog protein, e.g., aberrant activation of a hedgehog pathway. The loss-of-function may include a loss of the ability of the *ptc* gene product to regulate the level of expression of Ci genes, e.g., *Gli1*, *Gli2* and *Gli3*.

A "patient" or "subject" to be treated by the subject method can mean either a human or non-human animal.

The term "prodrug" is intended to encompass compounds which, under physiological conditions, are converted into the therapeutically active agents of the present invention. A common method for making a prodrug is to include selected moieties which are hydrolyzed under physiological conditions to reveal the desired molecule. In other embodiments, the prodrug is converted by an enzymatic activity of the host animal.

As used herein, "proliferating" and "proliferation" refer to cells undergoing mitosis.

Throughout this application, the term "proliferative skin disorder" refers to any disease/disorder of the skin marked by unwanted or aberrant proliferation of cutaneous tissue. These conditions are typically characterized by epidermal cell proliferation or incomplete cell differentiation, and include, for example, X-linked ichthyosis, psoriasis, atopic dermatitis, allergic contact dermatitis, epidermolytic hyperkeratosis, and seborrheic dermatitis. For example, epidermodysplasia is a form of faulty development of the epidermis. Another example is "epidermolysis", which refers to a loosened state of the epidermis with formation of blebs and bullae either spontaneously or at the site of trauma.

As used herein, the term "psoriasis" refers to a hyperproliferative skin disorder which alters the skin's regulatory mechanisms. In particular, lesions are formed which involve primary and secondary alterations in epidermal proliferation, inflammatory responses of the skin, and an expression of regulatory molecules such as lymphokines and inflammatory factors. Psoriatic skin is morphologically characterized by an increased turnover of epidermal cells, thickened epidermis, abnormal keratinization, inflammatory cell infiltrates into the dermis layer and polymorphonuclear leukocyte infiltration into the epidermis layer resulting in an increase in the basal cell cycle. Additionally, hyperkeratotic and parakeratotic cells are present.

The term "skin" refers to the outer protective covering of the body, consisting of the corium and the epidermis, and is understood to include sweat and sebaceous glands, as well as hair follicle structures. Throughout the present application, the adjective "cutaneous" may be used, and should be understood to refer generally to attributes of the skin, as appropriate to the context in which they are used.

The term "*smoothened* gain-of-function" refers to an aberrant modification or mutation of a *smo* gene, or an increased level of expression of the gene, which results in a phenotype which resembles contacting a cell with a hedgehog protein, e.g., aberrant activation of a hedgehog pathway. While not wishing to be bound by any particular theory, it is noted that *ptc* may not signal directly into the cell, but rather interact with *smoothened*, another membrane-bound protein located downstream of *ptc* in *hedgehog* signaling (Marigo et al., (1996) Nature 384: 177-179). The gene *smo* is a segment-

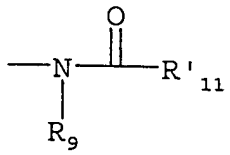
polarity gene required for the correct patterning of every segment in Drosophila (Alcedo et al., (1996) Cell 86: 221-232). Human homologs of *smo* have been identified. See, for example, Stone et al. (1996) Nature 384:129-134, and GenBank accession U84401. The *smoothened* gene encodes an integral membrane protein with characteristics of heterotrimeric G-protein-coupled receptors; i.e., 7-transmembrane regions. This protein shows homology to the Drosophila *Frizzled* (Fz) protein, a member of the *wingless* pathway. It was originally thought that *smo* encodes a receptor of the Hh signal. However, this suggestion was subsequently disproved, as evidence for *ptc* being the Hh receptor was obtained. Cells that express *Smo* fail to bind Hh, indicating that *smo* does not interact directly with Hh (Nusse, (1996) Nature 384: 119-120). Rather, the binding of *Sonic hedgehog* (SHH) to its receptor, *PTC*H, is thought to prevent normal inhibition by *PTC*H of *smoothened* (SMO), a seven-span transmembrane protein.

Recently, it has been reported that activating *smoothened* mutations occur in sporadic basal cell carcinoma, Xie et al. (1998) Nature 391: 90-2, and primitive neuroectodermal tumors of the central nervous system, Reifenberger et al. (1998) Cancer Res 58: 1798-803.

The term "therapeutic index" refers to the therapeutic index of a drug defined as LD_{50}/ED_{50}.

As used herein, "transformed cells" refers to cells which have spontaneously converted to a state of unrestrained growth, i.e., they have acquired the ability to grow through an indefinite number of divisions in culture. Transformed cells may be characterized by such terms as neoplastic, anaplastic and/or hyperplastic, with respect to their loss of growth control.

The term "acylamino" is art-recognized and refers to a moiety that can be represented by the general formula:



wherein R_9 is as defined above, and R'_{11} represents a hydrogen, an alkyl, an alkenyl or $-(CH_2)_m-R_8$, where m and R_8 are as defined above.

Herein, the term "aliphatic group" refers to a straight-chain, branched-chain, or cyclic aliphatic hydrocarbon group and includes saturated and unsaturated aliphatic groups, such as an alkyl group, an alkenyl group, and an alkynyl group.

The terms "alkenyl" and "alkynyl" refer to unsaturated aliphatic groups analogous in length and possible substitution to the alkyls described above, but that contain at least one double or triple bond respectively.

The terms "alkoxyl" or "alkoxy" as used herein refers to an alkyl group, as defined above, having an oxygen radical attached thereto. Representative alkoxyl groups include methoxy, ethoxy, propyloxy, tert-butoxy and the like. An "ether" is two hydrocarbons covalently linked by an oxygen. Accordingly, the substituent of an alkyl that renders that alkyl an ether is or resembles an alkoxyl, such as can be represented by one of -O-alkyl, -O-alkenyl, -O-alkynyl, $-O-(CH_2)_m-R_8$, where m and R_8 are described above.

The term "alkyl" refers to the radical of saturated aliphatic groups, including straight-chain alkyl groups, branched-chain alkyl groups, cycloalkyl (alicyclic) groups, alkyl-substituted cycloalkyl groups, and cycloalkyl-substituted alkyl groups. In preferred embodiments, a straight chain or branched chain alkyl has 30 or fewer carbon atoms in its backbone (e.g., C_1-C_{30} for straight chains, C_3-C_{30} for branched chains), and more preferably 20 or fewer. Likewise, preferred cycloalkyls have from 3-10 carbon atoms in their ring structure, and more preferably have 5, 6 or 7 carbons in the ring structure.

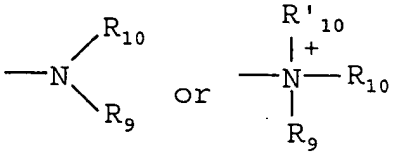
Moreover, the term "alkyl" (or "lower alkyl") as used throughout the specification, examples, and claims is intended to include both "unsubstituted alkyls" and "substituted alkyls", the latter of which refers to alkyl moieties having substituents replacing a hydrogen on one or more carbons of the hydrocarbon backbone. Such substituents can

include, for example, a halogen, a hydroxyl, a carbonyl (such as a carboxyl, an alkoxycarbonyl, a formyl, or an acyl), a thiocarbonyl (such as a thioester, a thioacetate, or a thioformate), an alkoxyl, a phosphoryl, a phosphate, a phosphonate, a phosphinate, an amino, an amido, an amidine, an imine, a cyano, a nitro, an azido, a sulfhydryl, an alkylthio, a sulfate, a sulfonate, a sulfamoyl, a sulfonamido, a sulfonyl, a heterocyclyl, an aralkyl, or an aromatic or heteroaromatic moiety. It will be understood by those skilled in the art that the moieties substituted on the hydrocarbon chain can themselves be substituted, if appropriate. For instance, the substituents of a substituted alkyl may include substituted and unsubstituted forms of amino, azido, imino, amido, phosphoryl (including phosphonate and phosphinate), sulfonyl (including sulfate, sulfonamido, sulfamoyl and sulfonate), and silyl groups, as well as ethers, alkylthios, carbonyls (including ketones, aldehydes, carboxylates, and esters), $-CF_3$, -CN and the like. Exemplary substituted alkyls are described below. Cycloalkyls can be further substituted with alkyls, alkenyls, alkoxys, alkylthios, aminoalkyls, carbonyl-substituted alkyls, $-CF_3$, -CN, and the like.

Unless the number of carbons is otherwise specified, "lower alkyl" as used herein means an alkyl group, as defined above, but having from one to ten carbons, more preferably from one to six carbon atoms in its backbone structure. Likewise, "lower alkenyl" and "lower alkynyl" have similar chain lengths. Throughout the application, preferred alkyl groups are lower alkyls. In preferred embodiments, a substituent designated herein as alkyl is a lower alkyl.

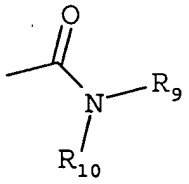
The term "alkylthio" refers to an alkyl group, as defined above, having a sulfur radical attached thereto. In preferred embodiments, the "alkylthio" moiety is represented by one of -S-alkyl, -S-alkenyl, -S-alkynyl, and $-S-(CH_2)_m-R_8$, wherein m and R_8 are defined above. Representative alkylthio groups include methylthio, ethylthio, and the like.

The terms "amine" and "amino" are art-recognized and refer to both unsubstituted and substituted amines, e.g., a moiety that can be represented by the general formula:



wherein R_9, R_{10} and R'_{10} each independently represent a hydrogen, an alkyl, an alkenyl, $-(CH_2)_m-R_8$, or R_9 and R_{10} taken together with the N atom to which they are attached complete a heterocycle having from 4 to 8 atoms in the ring structure; R_8 represents an aryl, a cycloalkyl, a cycloalkenyl, a heterocycle or a polycycle; and m is zero or an integer in the range of 1 to 8. In preferred embodiments, only one of R_9 or R_{10} can be a carbonyl, e.g., R_9, R_{10} and the nitrogen together do not form an imide. In even more preferred embodiments, R_9 and R_{10} (and optionally R'_{10}) each independently represent a hydrogen, an alkyl, an alkenyl, or $-(CH_2)_m-R_8$. Thus, the term "alkylamine" as used herein means an amine group, as defined above, having a substituted or unsubstituted alkyl attached thereto, i.e., at least one of R_9 and R_{10} is an alkyl group.

The term "amido" is art-recognized as an amino-substituted carbonyl and includes a moiety that can be represented by the general formula:



wherein R_9, R_{10} are as defined above. Preferred embodiments of the amide will not include imides which may be unstable.

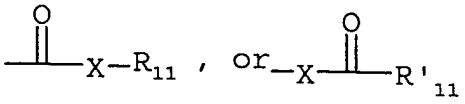
The term "aralkyl", as used herein, refers to an alkyl group substituted with an aryl group (e.g., an aromatic or heteroaromatic group).

The term "aryl" as used herein includes 5-, 6-, and 7-membered single-ring aromatic groups that may include from zero to four heteroatoms, for example, benzene, pyrrole, furan, thiophene, imidazole, oxazole, thiazole, triazole, pyrazole, pyridine, pyrazine, pyridazine and pyrimidine, and the like. Those aryl groups having heteroatoms in the ring structure may also be referred to as "aryl heterocycles" or "heteroaromatics."

The aromatic ring can be substituted at one or more ring positions with such substituents as described above, for example, halogen, azide, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, alkoxyl, amino, nitro, sulfhydryl, imino, amido, phosphate, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, sulfonamido, ketone, aldehyde, ester, heterocyclyl, aromatic or heteroaromatic moieties, $-CF_3$, -CN, or the like. The term "aryl" also includes polycyclic ring systems having two or more cyclic rings in which two or more carbons are common to two adjoining rings (the rings are "fused rings") wherein at least one of the rings is aromatic, e.g., the other cyclic rings can be cycloalkyls, cycloalkenyls, cycloalkynyls, aryls and/or heterocyclyls.

The term "carbocycle", as used herein, refers to an aromatic or non-aromatic ring in which each atom of the ring is carbon.

The term "carbonyl" is art-recognized and includes such moieties as can be represented by the general formula:



wherein X is a bond or represents an oxygen or a sulfur, and R_{11} represents a hydrogen, an alkyl, an alkenyl, $-(CH_2)_m-R_8$ or a pharmaceutically acceptable salt, R'_{11} represents a hydrogen, an alkyl, an alkenyl or $-(CH_2)_m-R_8$, where m and R_8 are as defined above. Where X is an oxygen and R_{11} or R'_{11} is not hydrogen, the formula represents an "ester". Where X is an oxygen, and R_{11} is as defined above, the moiety is referred to herein as a carboxyl group, and particularly when R_{11} is a hydrogen, the formula represents a "carboxylic acid". Where X is an oxygen, and R'_{11} is hydrogen, the formula represents a "formate". In general, where the oxygen atom of the above formula is replaced by sulfur, the formula represents a "thiocarbonyl" group. Where X is a sulfur and R_{11} or R'_{11} is not hydrogen, the formula represents a "thioester." Where X is a sulfur and R_{11} is hydrogen, the formula represents a "thiocarboxylic acid." Where X is a sulfur and R_{11}' is hydrogen, the formula represents a "thiolformate." On the other hand, where X is a bond, and R_{11} is

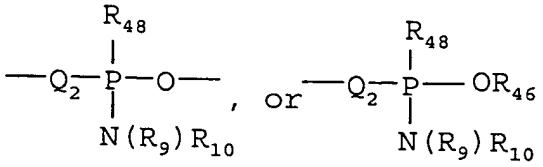
not hydrogen, the above formula represents a "ketone" group. Where X is a bond, and R_{11} is hydrogen, the above formula represents an "aldehyde" group.

The term "heteroatom" as used herein means an atom of any element other than carbon or hydrogen. Preferred heteroatoms are boron, nitrogen, oxygen, phosphorus, sulfur and selenium.

The terms "heterocyclyl" or "heterocyclic group" refer to 3- to 10-membered ring structures, more preferably 3- to 7-membered rings, whose ring structures include one to four heteroatoms. Heterocycles can also be polycycles. Heterocyclyl groups include, for example, thiophene, thianthrene, furan, pyran, isobenzofuran, chromene, xanthene, phenoxathiin, pyrrole, imidazole, pyrazole, isothiazole, isoxazole, pyridine, pyrazine, pyrimidine, pyridazine, indolizine, isoindole, indole, indazole, purine, quinolizine, isoquinoline, quinoline, phthalazine, naphthyridine, quinoxaline, quinazoline, cinnoline, pteridine, carbazole, carboline, phenanthridine, acridine, pyrimidine, phenanthroline, phenazine, phenarsazine, phenothiazine, furazan, phenoxazine, pyrrolidine, oxolane, thiolane, oxazole, piperidine, piperazine, morpholine, lactones, lactams such as azetidinones and pyrrolidinones, sultams, sultones, and the like. The heterocyclic ring can be substituted at one or more positions with such substituents as described above, as for example, halogen, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, amino, nitro, sulfhydryl, imino, amido, phosphate, phosphonate, phosphinate, carbonyl, carboxyl, silyl, ether, alkylthio, sulfonyl, ketone, aldehyde, ester, a heterocyclyl, an aromatic or heteroaromatic moiety, $-CF_3$, -CN, or the like.

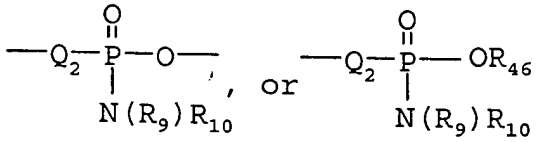
As used herein, the term "nitro" means $-NO_2$; the term "halogen" designates -F, -Cl, -Br or -I; the term "sulfhydryl" means -SH; the term "hydroxyl" means -OH; and the term "sulfonyl" means $-SO_2-$.

A "phosphonamidite" can be represented in the general formula:



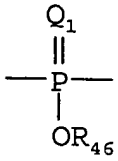
wherein R_9 and R_{10} are as defined above, Q_2 represents O, S or N, and R_{48} represents a lower alkyl or an aryl, Q_2 represents O, S or N.

A "phosphoramidite" can be represented in the general formula:

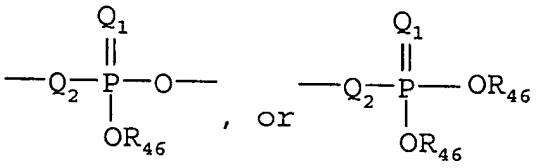


wherein R_9 and R_{10} are as defined above, and Q_2 represents O, S or N.

A "phosphoryl" can in general be represented by the formula:



wherein Q_1 represented S or O, and R_{46} represents hydrogen, a lower alkyl or an aryl. When used to substitute, for example, an alkyl, the phosphoryl group of the phosphorylalkyl can be represented by the general formula:



wherein Q_1 represented S or O, and each R_{46} independently represents hydrogen, a lower alkyl or an aryl, Q_2 represents O, S or N. When Q_1 is an S, the phosphoryl moiety is a "phosphorothioate".

The terms "polycyclyl" or "polycyclic group" refer to two or more rings (e.g., cycloalkyls, cycloalkenyls, cycloalkynyls, aryls and/or heterocyclyls) in which two or more carbons are common to two adjoining rings, e.g., the rings are "fused rings". Rings that are joined through non-adjacent atoms are termed "bridged" rings. Each of the rings of the polycycle can be substituted with such substituents as described above, as for example, halogen, alkyl, aralkyl, alkenyl, alkynyl, cycloalkyl, hydroxyl, amino, nitro, sulfhydryl, imino, amido, phosphate, phosphonate, phosphinate, carbonyl, carboxyl, silyl,

ether, alkylthio, sulfonyl, ketone, aldehyde, ester, a heterocyclyl, an aromatic or heteroaromatic moiety, $-CF_3$, -CN, or the like.

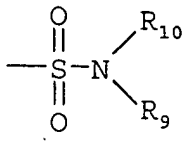
The phrase "protecting group" as used herein means temporary substituents which protect a potentially reactive functional group from undesired chemical transformations. Examples of such protecting groups include esters of carboxylic acids, silyl ethers of alcohols, and acetals and ketals of aldehydes and ketones, respectively. The field of protecting group chemistry has been reviewed (Greene, T.W.; Wuts, P.G.M. *Protective Groups in Organic Synthesis*, 2nd ed.; Wiley: New York, 1991).

A "selenoalkyl" refers to an alkyl group having a substituted seleno group attached thereto. Exemplary "selenoethers" which may be substituted on the alkyl are selected from one of -Se-alkyl, -Se-alkenyl, -Se-alkynyl, and $-Se-(CH_2)_m-R_8$, m and R_8 being defined above.

As used herein, the term "substituted" is contemplated to include all permissible substituents of organic compounds. In a broad aspect, the permissible substituents include acyclic and cyclic, branched and unbranched, carbocyclic and heterocyclic, aromatic and nonaromatic substituents of organic compounds. Illustrative substituents include, for example, those described herein above. The permissible substituents can be one or more and the same or different for appropriate organic compounds. For purposes of this invention, the heteroatoms such as nitrogen may have hydrogen substituents and/or any permissible substituents of organic compounds described herein which satisfy the valences of the heteroatoms. This invention is not intended to be limited in any manner by the permissible substituents of organic compounds.

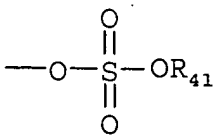
It will be understood that "substitution" or "substituted with" includes the implicit proviso that such substitution is in accordance with permitted valence of the substituted atom and the substituent, and that the substitution results in a stable compound, e.g., which does not spontaneously undergo transformation such as by rearrangement, cyclization, elimination, etc.

The term "sulfamoyl" is art-recognized and includes a moiety that can be represented by the general formula:



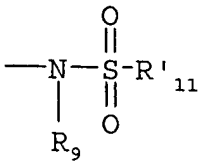
in which R_9 and R_{10} are as defined above.

The term "sulfate" is art recognized and includes a moiety that can be represented by the general formula:



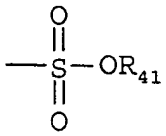
in which R_{41} is as defined above.

The term "sulfonamido" is art recognized and includes a moiety that can be represented by the general formula:



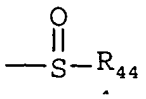
in which R_9 and R'_{11} are as defined above.

The term "sulfonate" is art-recognized and includes a moiety that can be represented by the general formula:



in which R_{41} is an electron pair, hydrogen, alkyl, cycloalkyl, or aryl.

The terms "sulfoxido" or "sulfinyl", as used herein, refers to a moiety that can be represented by the general formula:



in which R_{44} is selected from the group consisting of hydrogen, alkyl, alkenyl, alkynyl, cycloalkyl, heterocyclyl, aralkyl, or aryl.

Analogous substitutions can be made to alkenyl and alkynyl groups to produce, for example, aminoalkenyls, aminoalkynyls, amidoalkenyls, amidoalkynyls, iminoalkenyls, iminoalkynyls, thioalkenyls, thioalkynyls, carbonyl-substituted alkenyls or alkynyls.

As used herein, the definition of each expression, e.g., alkyl, m, n, etc., when it occurs more than once in any structure, is intended to be independent of its definition elsewhere in the same structure.

The terms triflyl, tosyl, mesyl, and nonaflyl are art-recognized and refer to trifluoromethanesulfonyl, *p*-toluenesulfonyl, methanesulfonyl, and nonafluorobutanesulfonyl groups, respectively. The terms triflate, tosylate, mesylate, and nonaflate are art-recognized and refer to trifluoromethanesulfonate ester, *p*-toluenesulfonate ester, methanesulfonate ester, and nonafluorobutanesulfonate ester functional groups and molecules that contain said groups, respectively.

The abbreviations Me, Et, Ph, Tf, Nf, Ts, Ms represent methyl, ethyl, phenyl, trifluoromethanesulfonyl, nonafluorobutanesulfonyl, *p*-toluenesulfonyl and methanesulfonyl, respectively. A more comprehensive list of the abbreviations utilized by organic chemists of ordinary skill in the art appears in the first issue of each volume of the *Journal of Organic Chemistry*; this list is typically presented in a table entitled Standard List of Abbreviations. The abbreviations contained in said list, and all abbreviations utilized by organic chemists of ordinary skill in the art are hereby incorporated by reference.

Certain compounds of the present invention may exist in particular geometric or stereoisomeric forms. The present invention contemplates all such compounds, including cis- and trans-isomers, *R*- and *S*-enantiomers, diastereomers, (D)-isomers, (L)-isomers, the racemic mixtures thereof, and other mixtures thereof, as falling within the scope of the invention. Additional asymmetric carbon atoms may be present in a substituent such as an

alkyl group. All such isomers, as well as mixtures thereof, are intended to be included in this invention.

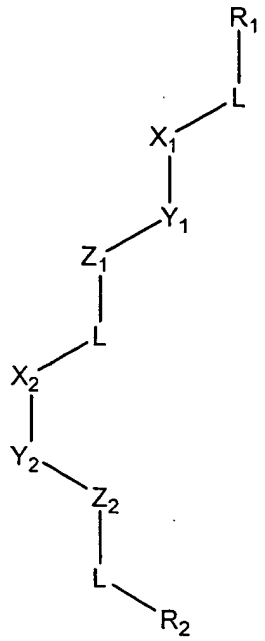
If, for instance, a particular enantiomer of a compound of the present invention is desired, it may be prepared by asymmetric synthesis, or by derivation with a chiral auxiliary, where the resulting diastereomeric mixture is separated and the auxiliary group cleaved to provide the pure desired enantiomers. Alternatively, where the molecule contains a basic functional group, such as amino, or an acidic functional group, such as carboxyl, diastereomeric salts may be formed with an appropriate optically active acid or base, followed by resolution of the diastereomers thus formed by fractional crystallization or chromatographic means well known in the art, and subsequent recovery of the pure enantiomers.

Contemplated equivalents of the compounds described above include compounds which otherwise correspond thereto, and which have the same general properties thereof (e.g., the ability to inhibit hedgehog signaling), wherein one or more simple variations of substituents are made which do not adversely affect the efficacy of the compound. In general, the compounds of the present invention may be prepared by the methods illustrated in the general reaction schemes as, for example, described below, or by modifications thereof, using readily available starting materials, reagents and conventional synthesis procedures. In these reactions, it is also possible to make use of variants which are in themselves known, but are not mentioned here.

For purposes of this invention, the chemical elements are identified in accordance with the Periodic Table of the Elements, CAS version, Handbook of Chemistry and Physics, 67th Ed., 1986-87, inside cover. Also for purposes of this invention, the term "hydrocarbon" is contemplated to include all permissible compounds having at least one hydrogen and one carbon atom. In a broad aspect, the permissible hydrocarbons include acyclic and cyclic, branched and unbranched, carbocyclic and heterocyclic, aromatic and nonaromatic organic compounds which can be substituted or unsubstituted.

III. Exemplary Compounds of the Invention.

As described in further detail below, it is contemplated that the subject methods can be carried out using a variety of different small molecules which can be readily identified, for example, by such drug screening assays as described herein. For example, compounds useful in the subject methods include compounds may be represented by general forumla (I):



Formula I

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, aryl (e.g., substituted or unsubstituted), aralkyl (e.g., substituted or unsubstituted, e.g., $-(CH_2)_n$aryl), or heteroaryl (e.g., substituted or unsubstituted), or heteroaralkyl (e.g., substituted or unsubstituted, e.g., $-(CH_2)_n$heteroaralkyl-);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

X_1 and X_2 can be selected, independently, from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-N(R_8)-, -ON(R_8)-, a heterocycle, or a direct bond between L and Y_1 or Y_2, respectively;

Y_1 and Y_2 can be selected, independently, from -C(=O)-, -C(=S)-, -S(O_2)-, -S(O)-, -C(=NCN)-, -P(=O)(OR_2)-, a heteroaromatic group, or a direct bond between X_1 and Z_1 or X_2 and Z_2, respectively;

Z_1 and Z_2 can be selected, independently, from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y_1 or Y_2, respectively, and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted), or two R_8 taken together may form a 4- to 8-membered ring, e.g., with X_1 and Z_1 or X_2 and Z_1, which ring may include one or more carbonyls;

p represents, independently for each occurrence, an integer from 0 to 10, preferably from 0 to 3; and

n, individually for each occurence, represents an integer from 0 to 10, preferably from 0 to 5.

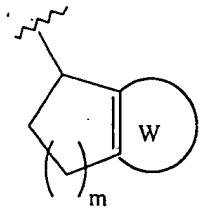
In certain embodiments, R_1 represents a substituted or unsubstituted heteroaryl group.

In certain embodiments, X_1 and X_2 can be selected from -N(R_8)-, -O-, -S-, a direct bond, and a heterocycle, Y_1 and Y_2 can be selected from -C(=O)-, -C(=S)-, and -S(O_2)-, and Z_1 or Z_2 can be selected from -N(R_8)-, -O-, -S-, a direct bond, and a heterocycle.

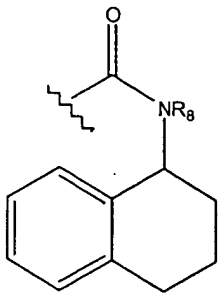
In certain related embodiments, X_1-Y_1-Z_1 or X_2-Y_2-Z_2 taken together represents a urea (N-C(O)-N) or an amide (N-C(O) or C(O)-N).

In certain embodiments, X_1 or X_2 represents a diazacarbocycle, such as a piperazine.

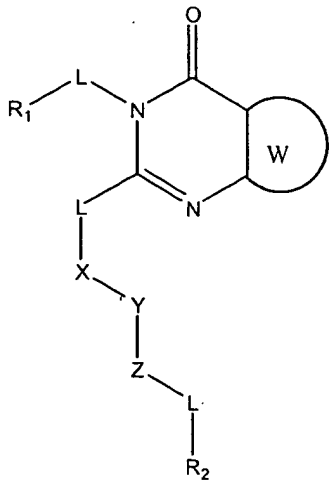
In certain embodiments, R_1 represents a fused cycloalkyl-aryl or cycloalkyl-heteroaryl system, for example:



wherein W is a substituted or unsubstituted aryl or heteroaryl ring fused to the cycloalkyl ring and m is an integer from 1-4 inclusive, e.g., from 1-3, or from 1-2. The fused system may be bound to L from any carbon of the fused system, including the position depicted above. In certain embodiments, R_1 may represent a tetrahydronaphthyl group, and preferably Y_1-X_1-L-R_1 taken together represent a tetrahydronaphthyl amide group, such as:



In embodiments wherein Y_1 and Z_1 are absent and X_1 comprises a pyrimidone, compounds useful in the present invention may be represented by general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X can be selected from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-$N(R_8)$-, -$ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y can be selected from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -$P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z can be selected from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted),

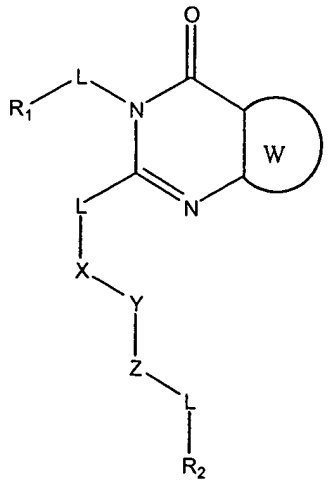
or two R_8 taken together may form a 4- to 8-membered ring, e.g., with X and Z, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10, preferably from 0 to 3; and

n, individually for each occurence, represents an integer from 0 to 10, preferably from 0 to 5.

In embodiments wherein Y_1 and Z_1 are absent and X_1 comprises a pyrimidone, compounds useful in the present invention may be represented by general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, aryl (e.g., substituted or unsubstituted), aralkyl (e.g., substituted or unsubstituted, e.g., $-(CH_2)_n$aryl), or heteroaryl (e.g., substituted or unsubstituted), or heteroaralkyl (e.g., substituted or unsubstituted, e.g., $-(CH_2)_n$heteroaralkyl-);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, -

$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-, which may optionally be substitued with a group selected from H, substituted or unsubstituted lower alkyl, alkenyl, or alkynyl, cycloalkylalkyl (e.g., substituted or unsubstituted, e.g., -$(CH_2)_n$cycloalkyl), (e.g., substituted or unsubstituted), aryl (e.g., substituted or unsubstituted), aralkyl (e.g., substituted or unsubstituted, e.g., -$(CH_2)_n$aryl), or heteroaryl (e.g., substituted or unsubstituted), or heteroaralkyl (e.g., substituted or unsubstituted, e.g., -$(CH_2)_n$heteroaralkyl-), preferably from H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted);

X can be selected from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-$N(R_8)$-, -$ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y can be selected from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -$P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z can be selected from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, aryl (e.g., substituted or unsubstituted), aralkyl (e.g., substituted or unsubstituted, e.g., -$(CH_2)_n$aryl), or heteroaryl (e.g., substituted or unsubstituted), or heteroaralkyl (e.g., substituted or unsubstituted, e.g., -$(CH_2)_n$heteroaralkyl-), or two R_8 taken together may form a 4- to 8-membered ring, e.g., with X and Z, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10, preferably from 0 to 3; and

n, individually for each occurrence, represents an integer from 0 to 10, preferably from 0 to 5.

In certain embodiments, R_1 represents a substituted or unsubstituted aryl or heteroaryl group, e.g., a phenyl ring, a pyridine ring, etc. In certain embodiments wherein $-LR_1$ represents a substituted aryl or heteroaryl group, R_1 is preferably not substituted with an isopropoxy (Me_2CHO-) group. In certain embodiments wherein $-LR_1$ represents a substituted aryl or heteroaryl group, R_1 is preferably not substituted with an ether group. In certain embodiments, substituents on R_1 (e.g., other than hydrogen) are selected from halogen, cyano, alkyl, alkenyl, alkynyl, aryl, hydroxyl, (unbranched alkyl-O-), silyloxy, amino, nitro, thiol, amino, imino, amido, phosphoryl, phosphonate, phosphine, carbonyl, carboxyl, carboxamide, anhydride, silyl, thioether, alkylsulfonyl, arylsulfonyl, sulfoxide, selenoether, ketone, aldehyde, ester, or $-(CH_2)_m-R_8$. In certain embodiments, non-hydrogen substituents are selected from halogen, cyano, alkyl, alkenyl, alkynyl, aryl, nitro, thiol, imino, amido, carbonyl, carboxyl, anhydride, thioether, alkylsulfonyl, arylsulfonyl, ketone, aldehyde, and ester. In certain embodiments, non-hydrogen substituents are selected from halogen, cyano, alkyl, alkenyl, alkynyl, nitro, amido, carboxyl, anhydride, alkylsulfonyl, ketone, aldehyde, and ester.

In certain embodiments, X can be selected from $-N(R_8)-$, -O-, -S-, a direct bond, and a heterocycle, Y can be selected from -C(=O)-, -C(=S)-, and $-S(O_2)-$, and Z can be selected from $-N(R_8)-$, -O-, -S-, a direct bond, and a heterocycle. In certain such embodiments, at least one of Z and X is present.

In certain related embodiments, X-Y-Z taken together represents a urea (NC(O)N) or an amide (NC(O) or C(O)N).

In certain embodiments, W is a substituted or unsubstituted benzene ring.

In certain embodiments, X represents a diazacarbocycle, such as a piperazine, e.g., substituted or unsubstituted.

In certain embodiments, X can be selected from $-N(R_8)-$, -O-, -S-, and a direct bond, Y can be selected from -C(=O)-, -C(=S)-, and $-S(O_2)-$, and Z can be selected from $-N(R_8)-$, -O-, -S-, and a direct bond, such that at least one of X and Z is present.

In certain embodiments R_8 represents H, lower alkyl, aralkyl, heteroaralkyl, aryl, or heteroaryl, e.g., H or lower alkyl.

In certain embodiments, X represents -NH-.

In certain embodiments, -L-X- represents -(unbranched lower alkyl)-NH-, e.g., -CH_2-NH-, -CH_2CH_2-NH-, etc.

In certain embodiments, the subject antagonists can be chosen on the basis of their selectively for the *hedgehog* pathway. This selectivity can be for the *hedgehog* pathway versus other pathways, or for selectivity between particular *hedgehog* pathways, e.g., e.g., *ptc*-1, *ptc*-2, etc.

In certain preferred embodiments, the subject inhibitors inhibit *hedgehog*-mediated signal transduction with an ED_{50} of 1 mM or less, more preferably of 1 μM or less, and even more preferably of 1 nM or less.

In particular embodiments, the small molecule is chosen for use because it is more selective for one patched isoform over the next, e.g., 10 fold, and more preferably at least 100 or even 1000 fold more selective for one patched pathway (*ptc*-1, *ptc*-2) over another.

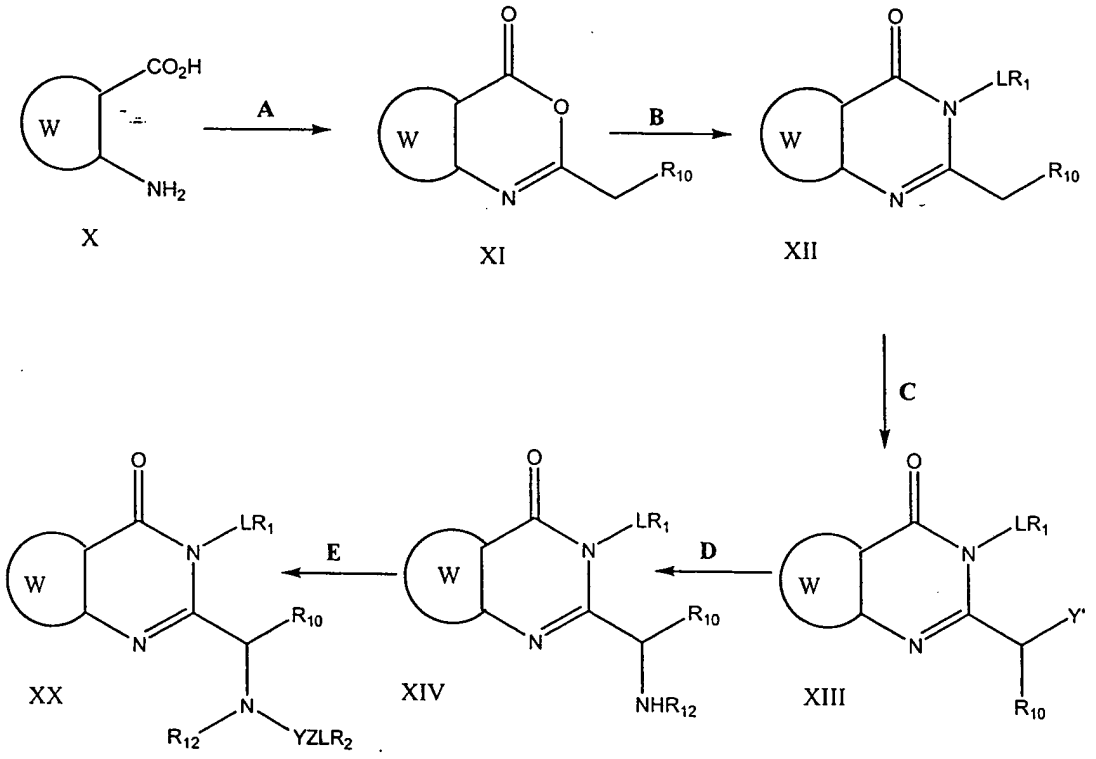
In certain embodiments, a compound which is an antagonist of the *hedgehog* pathway is chosen to selectively antagonize *hedgehog* activity over protein kinases other than PKA, such as PKC, e.g., the compound modulates the activity of the *hedgehog* pathway at least an order of magnitude more strongly than it modulates the activity of another protein kinase, preferably at least two orders of magnitude more strongly, even more preferably at least three orders of magnitude more strongly. Thus, for example, a preferred inhibitor of the *hedgehog* pathway may inhibit *hedgehog* activity with a K_i at least an order of magnitude lower than its K_i for inhibition of PKC, preferably at least two orders of magnitude lower, even more preferably at least three orders of magnitude lower. In certain embodiments, the K_i for PKA inhibition is less than 10 nM, preferably less than 1 nM, even more preferably less than 0.1 nM.

Synthesis of Subject Compounds

In certain embodiments, the present invention relates to techniques and approaches for the synthesis of subject compounds as described by Formulae I-VIII.

For example, in certain embodiments, the present invention relates to a method for preparing a subject compound as set forth in Scheme I,

Scheme I



wherein step (A) comprises reacting a compound having a structure of Formula X, wherein W represents a substituted or unsubstituted aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid group in adjacent (*ortho*) positions, with an acylating agent having the formula $R_{10}CH_2C(=O)X'$, wherein R_{10}, independently for each occurrence, represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl, and X' represents a halogen or $-OC(=O)CH_2R_{10}$, under conditions that produce a compound having a structure of Formula XI;

step (B) comprises reacting a compound having a structure of Formula XI with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XII;

step (C) comprises reacting a compound having a structure of Formula XII with a halogenating agent, such as chlorine, bromine, iodine, N-bromosuccinimide, N-chlorosuccinimide, N-iodosuccinimide, ClBr, IBr, ClI, or a reagent that generates a halogen radical (such as Cl· , Br· , or I·) under conditions that result in a compound having a structure of Formula XIII, wherein Y' represents a halogen such as Cl, Br, or I;

step (D) comprises reacting a compound having a structure of Formula XIII with an amine having the formula H_2NR_{12}, wherein R_{12} represents a lower alkyl group or a silyl group, such as a trialkylsilyl, triarylsilyl, dialkylarylsilyl, or diarylalkylsilyl group, under conditions that result in a compound having a structure of Formula XIV; and

step (E) comprises reacting a compound having a structure of Formula XIV with a terminating group having a structure of R_2V' to produce a compound having a structure of Formula XX, wherein R_2 is as defined above, and V' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, $ZC(=W)WC(=W)ZR_2$, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an activated acylating moiety prepared *in situ*.

In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XI and performing steps (B) to (E). In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XII and performing steps (C) to (E). In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XIII and performing steps (D) to (E). In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XIV and performing step (E).

In certain embodiments, step (A) may be performed using an anhydride, such as a symmetrical anhydride, having the formula $R_{10}C(=O)OC(=O)R_{10}$. In certain embodiments, the reaction may be performed using the anhydride as a solvent, e.g., the

reaction mixture is substantially free of a solvent other than the anhydride. In certain embodiments, R_{10}, for both occurrences, is a lower alkyl group.

In certain embodiments, step (B) may be performed by treating a compound of Formula XI with an amine in a non-polar solvent, such as chloroform, methylene chloride, ethylene chloride, toluene, benzene, ether, tetrahydrofuran, or another non-polar solvent, or any combination thereof, followed by treating the product of that reaction with an ionic base, such as an alkali metal hydroxide or alkoxide, such as methoxide, ethoxide, etc., in a polar solvent, such as an alcoholic solvent, e.g., methanol, ethanol, propanol, ethylene glycol, propylene glycol, glycerol, etc. In certain embodiments, the conditions may include heating the reaction mixture above room temperature, e.g., to at least 50 °C, or at least 75 °C, or even at least 100 °C. In certain embodiments, this step may convert a compound of Formula X to a compound of Formula XI in at least about 80% yield, or at least about 85% yield.

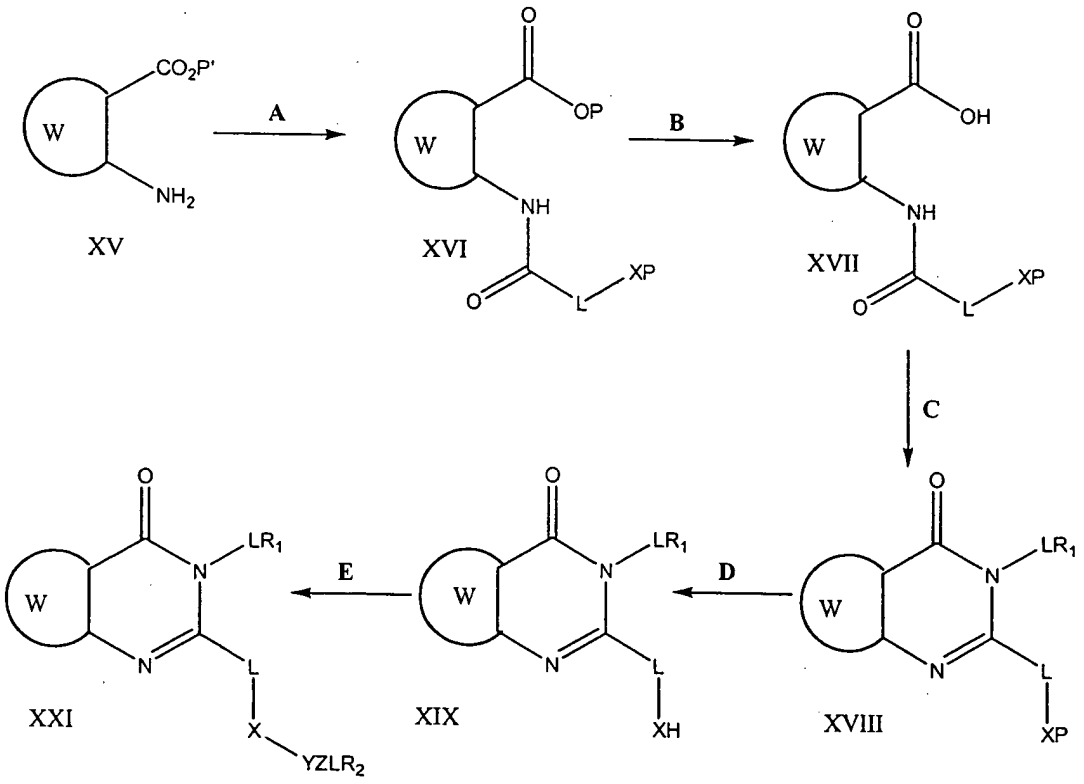
In certain embodiments, step (C) may be performed using a halogen as a halogenating agent, and may be performed in a reaction mixture including a carboxylic acid, such as acetic acid or propionic acid. In certain embodiments, a mild base, such as a carboxylic acid salt, may be added to or present in the reaction mixture.

In certain embodiments, step (D) may be performed using a polar solvent, such as water, ethanol, methanol, ethylene glycol, DMF, or another polar solvent, or any combination of polar solvents. In certain embodiments, the solvent or solvent mixture comprises less than about 50% water, or is non-aqueous, i.e., is substantially free of water. In certain embodiments, the polar solvent comprises an alcohol, such as methanol, ethanol, propanol, isopropanol, butanol, isobutanol, t-butanol, sec-butanol, ethylene glycol, and 1,3-propanediol.

In certain embodiments, step (E) may be performed by generating an acylating agent *in situ*, for example, by reacting a carboxylic acid with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an

increased susceptibility, relative to the carboxylic acid, towards coupling with an amine. A wide variety of such reagents are well known in the art of organic synthesis, especially peptide coupling. Similarly, a primary amine can be treated with a phosgene equivalent, such as carbonyl diimidazole, phosgene, triphosgene, diphosgene, etc., or a thiophosgene equivalent, such as thiophosgene, thiocarbonyldiimidazole, etc., to generate an acylating agent (e.g., an isocyanate, isothiocyanate, chloroformamide, or chlorothioformamide, for example) capable of reacting with an amine to form a urea or thiourea, without necessitating isolation or purification of the acylating agent.

In certain embodiments, the present invention relates to a method for preparing a subject compound as set forth in Scheme II,



wherein step (A) comprises reacting a compound having a structure of Formula XV, wherein P' represents H or a protecting group, W represents a substituted or unsubstituted aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid or ester group in adjacent (*ortho*) positions, with an acylating agent having the

formula PXLC(=O)X', wherein X and L are as defined above, P represents a protecting group, and X' represents a halogen, –OC(=O)LXP, or a functional group generated by reacting a carboxyl group with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, carbonyldiimidazole, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an increased susceptibility, relative to the carboxylic acid, towards coupling with an amine, under conditions that produce a compound having a structure of Formula XVI;

step (B) comprises deprotecting the ester of a compound having a structure of Formula XVI to produce a carboxylic acid having a structure of Formula XVII, if necessary;

step (C) comprises reacting a compound having a structure of Formula XVII with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XVIII;

step (D) comprises removing the protecting group P from a compound having a structure of Formula XVIII to generate a compound having a structure of Formula XIX;

step (E) comprises reacting a compound having a structure of Formula XIX with a terminating group having a structure of R_2Y' to produce a compound having a structure of Formula XXI, wherein R_2 is as defined above, and Y' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, $ZC(=W)WC(=W)ZR_2$, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an active acylating moiety prepared *in situ*.

In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XVI and performing steps (B) to (E). In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XVII and performing steps (C) to (E). In certain embodiments, a subject compound may be prepared by providing a compound having a structure of Formula XVIII and performing steps (D) to (E). In certain embodiments, a subject

compound may be prepared by providing a compound having a structure of Formula XIX and performing step (E).

In certain embodiments, step (B) may be performed by treating a compound of Formula XVI with an acid in the presence of water, or with a hydroxide base (e.g., to hydrolyze or saponify an ester), by hydrogenolysis (e.g., to remove benzyl or allyl esters), in the presence of a mild base (e.g., to remove a fluorenylmethyl ester), by adding a Lewis acid (e.g., to remove a methoxymethyl ester), in the presence of fluoride ion (e.g., to remove a 2-silylethyl ester), or any other suitable means.

In certain embodiments, step (C) may be performed by activating the carboxyl group of the compound of Formula XVII with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, carbonyldiimidazole, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an increased susceptibility, relative to the carboxylic acid, towards coupling with an amine. A wide variety of such reagents are well known in the art of organic synthesis, especially peptide coupling. In certain embodiments, the conditions may include heating the reaction mixture above room temperature, e.g., to at least 50 °C, or at least 75 °C, or even at least 100 °C.

In certain embodiments, step (D) may be performed by hydrogenolysis (e.g., to remove an Alloc or CBz group), by using a base or a hydride reagent (e.g., to remove a trifluoroacetyl group), by using an acid, such as trifluoroacetic acid (e.g., to remove a BOC group), or by any other means suitable to deprotect the amine.

In certain embodiments, step (E) may be performed by generating an acylating agent *in situ*, for example, by reacting a carboxylic acid with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an increased susceptibility, relative to the carboxylic acid, towards coupling with an amine.

A wide variety of such reagents are well known in the art of organic synthesis, especially peptide coupling. Similarly, a primary amine can be treated with a phosgene equivalent, such as carbonyl diimidazole, phosgene, triphosgene, diphosgene, etc., or a thiophosgene equivalent, such as thiophosgene, thiocarbonyldiimidazole, etc., to generate an acylating agent (e.g., an isocyanate, isothiocyanate, chloroformamide, or chlorothioformamide, for example) capable of reacting with an amine to form a urea or thiourea, without necessitating isolation or purification of the acylating agent.

IV. Exemplary Applications of Method and Compositions

Another aspect of the present invention relates to a method of modulating a differentiated state, survival, and/or proliferation of a cell having a *hedgehog* gain-of-function phenotype by contacting the cells with a *hedgehog* antagonist according to the subject method and as the circumstances may warrant.

For instance, it is contemplated by the invention that, in light of the findings of an apparently broad involvement of *hedgehog*, *ptc*, and *smoothened* in the formation of ordered spatial arrangements of differentiated tissues in vertebrates, the subject method could be used as part of a process for generating and/or maintaining an array of different vertebrate tissue both *in vitro* and *in vivo*. The *hedgehog* antagonist, whether inductive or anti-inductive with respect proliferation or differentiation of a given tissue, can be, as appropriate, any of the preparations described above.

For example, the present method is applicable to cell culture techniques wherein, whether for genetic or biochemical reasons, the cells have a *hedgehog* gain-of-function phenotype. *In vitro* neuronal culture systems have proved to be fundamental and indispensable tools for the study of neural development, as well as the identification of neurotrophic factors such as nerve growth factor (NGF), ciliary trophic factors (CNTF), and brain derived neurotrophic factor (BDNF). One use of the present method may be in cultures of neuronal stem cells, such as in the use of such cultures for the generation of new neurons and glia. In such embodiments of the subject method, the cultured cells can be contacted with a *hedgehog* antagonist of the present invention in order to alter the rate

of proliferation of neuronal stem cells in the culture and/or alter the rate of differentiation, or to maintain the integrity of a culture of certain terminally differentiated neuronal cells. In an exemplary embodiment, the subject method can be used to culture, for example, sensory neurons or, alternatively, motorneurons. Such neuronal cultures can be used as convenient assay systems as well as sources of implantable cells for therapeutic treatments.

According to the present invention, large numbers of non-tumorigenic neural progenitor cells can be perpetuated *in vitro* and their rate of proliferation and/or differentiation can be affected by contact with *hedgehog* antagonists of the present invention. Generally, a method is provided comprising the steps of isolating neural progenitor cells from an animal, perpetuating these cells *in vitro* or *in vivo*, preferably in the presence of growth factors, and regulating the differentiation of these cells into particular neural phenotypes, e.g., neurons and glia, by contacting the cells with a *hedgehog* antagonist.

Progenitor cells are thought to be under a tonic inhibitory influence which maintains the progenitors in a suppressed state until their differentiation is required. However, recent techniques have been provided which permit these cells to be proliferated, and unlike neurons which are terminally differentiated and therefore non-dividing, they can be produced in unlimited number and are highly suitable for transplantation into heterologous and autologous hosts with neurodegenerative diseases.

By "progenitor" it is meant an oligopotent or multipotent stem cell which is able to divide without limit and, under specific conditions, can produce daughter cells which terminally differentiate such as into neurons and glia. These cells can be used for transplantation into a heterologous or autologous host. By heterologous is meant a host other than the animal from which the progenitor cells were originally derived. By autologous is meant the identical host from which the cells were originally derived.

Cells can be obtained from embryonic, post-natal, juvenile or adult neural tissue from any animal. By any animal is meant any multicellular animal which contains

nervous tissue. More particularly, is meant any fish, reptile, bird, amphibian or mammal and the like. The most preferable donors are mammals, especially mice and humans.

In the case of a heterologous donor animal, the animal may be euthanized, and the brain and specific area of interest removed using a sterile procedure. Brain areas of particular interest include any area from which progenitor cells can be obtained which will serve to restore function to a degenerated area of the host's brain. These regions include areas of the central nervous system (CNS) including the cerebral cortex, cerebellum, midbrain, brainstem, spinal cord and ventricular tissue, and areas of the peripheral nervous system (PNS) including the carotid body and the adrenal medulla. More particularly, these areas include regions in the basal ganglia, preferably the striatum which consists of the caudate and putamen, or various cell groups such as the globus pallidus, the subthalamic nucleus, the nucleus basalis which is found to be degenerated in Alzheimer's Disease patients, or the substantia nigra pars compacta which is found to be degenerated in Parkinson's Disease patients.

Human heterologous neural progenitor cells may be derived from fetal tissue obtained from elective abortion, or from a post-natal, juvenile or adult organ donor. Autologous neural tissue can be obtained by biopsy, or from patients undergoing neurosurgery in which neural tissue is removed, in particular during epilepsy surgery, and more particularly during temporal lobectomies and hippocampalectomies.

Cells can be obtained from donor tissue by dissociation of individual cells from the connecting extracellular matrix of the tissue. Dissociation can be obtained using any known procedure, including treatment with enzymes such as trypsin, collagenase and the like, or by using physical methods of dissociation such as with a blunt instrument or by mincing with a scalpel to a allow outgrowth of specific cell types from a tissue. Dissociation of fetal cells can be carried out in tissue culture medium, while a preferable medium for dissociation of juvenile and adult cells is artificial cerebral spinal fluid (aCSF). Regular aCSF contains 124 mM NaCl, 5 mM KCl, 1.3 mM $MgCl_2$, 2 mM $CaCl_2$, 26 mM $NaHCO_3$, and 10 mM D-glucose. Low Ca^{2+} aCSF contains the same

ingredients except for $MgCl_2$ at a concentration of 3.2 mM and $CaCl_2$ at a concentration of 0.1 mM.

Dissociated cells can be placed into any known culture medium capable of supporting cell growth, including MEM, DMEM, RPMI, F-12, and the like, containing supplements which are required for cellular metabolism such as glutamine and other amino acids, vitamins, minerals and useful proteins such as transferrin and the like. Medium may also contain antibiotics to prevent contamination with yeast, bacteria and fungi such as penicillin, streptomycin, gentamicin and the like. In some cases, the medium may contain serum derived from bovine, equine, chicken and the like. A particularly preferable medium for cells is a mixture of DMEM and F-12.

Conditions for culturing should be close to physiological conditions. The pH of the culture media should be close to physiological pH, preferably between pH 6-8, more preferably close to pH 7, even more particularly about pH 7.4. Cells should be cultured at a temperature close to physiological temperature, preferably between 30 °C-40 °C, more preferably between 32 °C-38 °C, and most preferably between 35 °C-37 °C.

Cells can be grown in suspension or on a fixed substrate, but proliferation of the progenitors is preferably done in suspension to generate large numbers of cells by formation of "neurospheres" (see, for example, Reynolds et al. (1992) *Science* 255:1070-1709; and PCT Publications WO93/01275, WO94/09119, WO94/10292, and WO94/16718). In the case of propagating (or splitting) suspension cells, flasks are shaken well and the neurospheres allowed to settle on the bottom corner of the flask. The spheres are then transferred to a 50 ml centrifuge tube and centrifuged at low speed. The medium is aspirated, the cells resuspended in a small amount of medium with growth factor, and the cells mechanically dissociated and resuspended in separate aliquots of media.

Cell suspensions in culture medium are supplemented with any growth factor which allows for the proliferation of progenitor cells and seeded in any receptacle capable of sustaining cells, though as set out above, preferably in culture flasks or roller bottles. Cells typically proliferate within 3-4 days in a 37 °C incubator, and proliferation can be

reinitiated at any time after that by dissociation of the cells and resuspension in fresh medium containing growth factors.

In the absence of substrate, cells lift off the floor of the flask and continue to proliferate in suspension forming a hollow sphere of undifferentiated cells. After approximately 3-10 days *in vitro*, the proliferating clusters (neurospheres) are fed every 2-7 days, and more particularly every 2-4 days by gentle centrifugation and resuspension in medium containing growth factor.

After 6-7 days *in vitro*, individual cells in the neurospheres can be separated by physical dissociation of the neurospheres with a blunt instrument, more particularly by triturating the neurospheres with a pipette. Single cells from the dissociated neurospheres are suspended in culture medium containing growth factors, and differentiation of the cells can be control in culture by plating (or resuspending) the cells in the presence of a *hedgehog* antagonist.

To further illustrate other uses of the subject *hedgehog* antagonists, it is noted that intracerebral grafting has emerged as an additional approach to central nervous system therapies. For example, one approach to repairing damaged brain tissues involves the transplantation of cells from fetal or neonatal animals into the adult brain (Dunnett et al. (1987) *J Exp Biol* 123:265-289; and Freund et al. (1985) *J Neurosci* 5:603-616). Fetal neurons from a variety of brain regions can be successfully incorporated into the adult brain, and such grafts can alleviate behavioral defects. For example, movement disorder induced by lesions of dopaminergic projections to the basal ganglia can be prevented by grafts of embryonic dopaminergic neurons. Complex cognitive functions that are impaired after lesions of the neocortex can also be partially restored by grafts of embryonic cortical cells. The subject method can be used to regulate the growth state in the culture, or where fetal tissue is used, especially neuronal stem cells, can be used to regulate the rate of differentiation of the stem cells.

Stem cells useful in the present invention are generally known. For example, several neural crest cells have been identified, some of which are multipotent and likely represent uncommitted neural crest cells, and others of which can generate only one type

of cell, such as sensory neurons, and likely represent committed progenitor cells. The role of *hedgehog* antagonists employed in the present method to culture such stem cells can be to regulate differentiation of the uncommitted progenitor, or to regulate further restriction of the developmental fate of a committed progenitor cell towards becoming a terminally differentiated neuronal cell. For example, the present method can be used *in vitro* to regulate the differentiation of neural crest cells into glial cells, schwann cells, chromaffin cells, cholinergic sympathetic or parasympathetic neurons, as well as peptidergic and serotonergic neurons. The *hedgehog* antagonists can be used alone, or can be used in combination with other neurotrophic factors which act to more particularly enhance a particular differentiation fate of the neuronal progenitor cell.

In addition to the implantation of cells cultured in the presence of the subject *hedgehog* antagonists, yet another aspect of the present invention concerns the therapeutic application of a *hedgehog* antagonist to regulate the growth state of neurons and other neuronal cells in both the central nervous system and the peripheral nervous system. The ability of *ptc*, *hedgehog*, and *smoothened* to regulate neuronal differentiation during development of the nervous system and also presumably in the adult state indicates that, in certain instances, the subject *hedgehog* antagonists can be expected to facilitate control of adult neurons with regard to maintenance, functional performance, and aging of normal cells; repair and regeneration processes in chemically or mechanically lesioned cells; and treatment of degeneration in certain pathological conditions. In light of this understanding, the present invention specifically contemplates applications of the subject method to the treatment protocol of (prevention and/or reduction of the severity of) neurological conditions deriving from: (i) acute, subacute, or chronic injury to the nervous system, including traumatic injury, chemical injury, vascular injury and deficits (such as the ischemia resulting from stroke), together with infectious/inflammatory and tumor-induced injury; (ii) aging of the nervous system including Alzheimer's disease; (iii) chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amylotrophic lateral sclerosis and the like, as well as spinocerebellar degenerations; and (iv) chronic immunological diseases of the nervous system or affecting the nervous system, including multiple sclerosis.

As appropriate, the subject method can also be used in generating nerve prostheses for the repair of central and peripheral nerve damage. In particular, where a crushed or severed axon is intubulated by use of a prosthetic device, *hedgehog* antagonists can be added to the prosthetic device to regulate the rate of growth and regeneration of the dendridic processes. Exemplary nerve guidance channels are described in U.S. patents 5,092,871 and 4,955,892.

In another embodiment, the subject method can be used in the treatment of neoplastic or hyperplastic transformations such as may occur in the central nervous system. For instance, the *hedgehog* antagonists can be utilized to cause such transformed cells to become either post-mitotic or apoptotic. The present method may, therefore, be used as part of a treatment for, e.g., malignant gliomas, meningiomas, medulloblastomas, neuroectodermal tumors, and ependymomas.

In a preferred embodiment, the subject method can be used as part of a treatment regimen for malignant medulloblastoma and other primary CNS malignant neuroectodermal tumors.

In certain embodiments, the subject method is used as part of treatment program for medulloblastoma. Medulloblastoma, a primary brain tumor, is the most common brain tumor in children. A medulloblastoma is a primitive neuroectodermal tumor arising in the posterior fossa. They account for approximately 25% of all pediatric brain tumors (Miller). Histologically, they are small round cell tumors commonly arranged in true rosettes, but may display some differentiation to astrocytes, ependymal cells or neurons (Rorke; Kleihues). PNET's may arise in other areas of the brain including the pineal gland (pineoblastoma) and cerebrum. Those arising in the supratentorial region generally fare worse than their PF counterparts.

Medulloblastoma/PNET's are known to recur anywhere in the CNS after resection, and can even metastasize to bone. Pretreatment evaluation should therefore include an examination of the spinal cord to exclude the possibility of "dropped metastases". Gadolinium-enhanced MRI has largely replaced myelography for this purpose, and CSF cytology is obtained postoperatively as a routine procedure.

In other embodiments, the subject method is used as part of treatment program for ependymomas. Ependymomas account for approximately 10% of the pediatric brain tumors in children. Grossly, they are tumors that arise from the ependymal lining of the ventricles and microscopically form rosettes, canals, and perivascular rosettes. In the CHOP series of 51 children reported with ependymomas, ¾ were histologically benign. Approximately 2/3 arose from the region of the 4th ventricle. One third presented in the supratentorial region. Age at presentation peaks between birth and 4 years, as demonstrated by SEER data as well as data from CHOP. The median age is about 5 years. Because so many children with this disease are babies, they often require multimodal therapy.

Yet another aspect of the present invention concerns the observation in the art that *ptc*, *hedgehog*, and/or *smoothened* are involved in morphogenic signals involved in other vertebrate organogenic pathways in addition to neuronal differentiation as described above, having apparent roles in other endodermal patterning, as well as both mesodermal and endodermal differentiation processes. Thus, it is contemplated by the invention that compositions comprising *hedgehog* antagonists can also be utilized for both cell culture and therapeutic methods involving generation and maintenance of non-neuronal tissue.

In one embodiment, the present invention makes use of the discovery that *ptc*, *hedgehog*, and *smoothened* are apparently involved in controlling the development of stem cells responsible for formation of the digestive tract, liver, lungs, and other organs which derive from the primitive gut. *Shh* serves as an inductive signal from the endoderm to the mesoderm, which is critical to gut morphogenesis. Therefore, for example, *hedgehog* antagonists of the instant method can be employed for regulating the development and maintenance of an artificial liver which can have multiple metabolic functions of a normal liver. In an exemplary embodiment, the subject method can be used to regulate the proliferation and differentiation of digestive tube stem cells to form hepatocyte cultures which can be used to populate extracellular matrices, or which can be encapsulated in biocompatible polymers, to form both implantable and extracorporeal artificial livers.

In another embodiment, therapeutic compositions of *hedgehog* antagonists can be utilized in conjunction with transplantation of such artificial livers, as well as embryonic liver structures, to regulate uptake of intraperitoneal implantation, vascularization, and *in vivo* differentiation and maintenance of the engrafted liver tissue.

In yet another embodiment, the subject method can be employed therapeutically to regulate such organs after physical, chemical or pathological insult. For instance, therapeutic compositions comprising *hedgehog* antagonists can be utilized in liver repair subsequent to a partial hepatectomy.

The generation of the pancreas and small intestine from the embryonic gut depends on intercellular signalling between the endodermal and mesodermal cells of the gut. In particular, the differentiation of intestinal mesoderm into smooth muscle has been suggested to depend on signals from adjacent endodermal cells. One candidate mediator of endodermally derived signals in the embryonic hindgut is Sonic hedgehog. See, for example, Apelqvist et al. (1997) Curr Biol 7:801-4. The Shh gene is expressed throughout the embryonic gut endoderm with the exception of the pancreatic bud endoderm, which instead expresses high levels of the homeodomain protein Ipf1/Pdx1 (insulin promoter factor 1/pancreatic and duodenal homeobox 1), an essential regulator of early pancreatic development. Apelqvist et al., supra, have examined whether the differential expression of Shh in the embryonic gut tube controls the differentiation of the surrounding mesoderm into specialised mesoderm derivatives of the small intestine and pancreas. To test this, they used the promoter of the Ipf1/Pdx1 gene to selectively express Shh in the developing pancreatic epithelium. In Ipf1/Pdx1- Shh transgenic mice, the pancreatic mesoderm developed into smooth muscle and interstitial cells of Cajal, characteristic of the intestine, rather than into pancreatic mesenchyme and spleen. Also, pancreatic explants exposed to Shh underwent a similar program of intestinal differentiation. These results provide evidence that the differential expression of endodermally derived Shh controls the fate of adjacent mesoderm at different regions of the gut tube.

In the context of the present invention, it is contemplated therefore that the subject *hedgehog* antagonists can be used to control or regulate the proliferation and/or differentiation of pancreatic tissue both *in vivo* and *in vitro*.

There are a wide variety of pathological cell proliferative and differentiative conditions for which the inhibitors of the present invention may provide therapeutic benefits, with the general strategy being, for example, the correction of aberrant insulin expression, or modulation of differentiation. More generally, however, the present invention relates to a method of inducing and/or maintaining a differentiated state, enhancing survival and/or affecting proliferation of pancreatic cells, by contacting the cells with the subject inhibitors. For instance, it is contemplated by the invention that, in light of the apparent involvement of *ptc*, *hedgehog*, and *smoothened* in the formation of ordered spatial arrangements of pancreatic tissues, the subject method could be used as part of a technique to generate and/or maintain such tissue both *in vitro* and *in vivo*. For instance, modulation of the function of *hedgehog* can be employed in both cell culture and therapeutic methods involving generation and maintenance β-cells and possibly also for non-pancreatic tissue, such as in controlling the development and maintenance of tissue from the digestive tract, spleen, lungs, urogenital organs (e.g., bladder), and other organs which derive from the primitive gut.

In an exemplary embodiment, the present method can be used in the treatment of hyperplastic and neoplastic disorders effecting pancreatic tissue, particularly those characterized by aberrant proliferation of pancreatic cells. For instance, pancreatic cancers are marked by abnormal proliferation of pancreatic cells which can result in alterations of insulin secretory capacity of the pancreas. For instance, certain pancreatic hyperplasias, such as pancreatic carcinomas, can result in hypoinsulinemia due to dysfunction of β-cells or decreased islet cell mass. To the extent that aberrant *ptc*, *hedgehog*, and *smoothened* signaling may be indicated in disease progression, the subject inhibitors, can be used to enhance regeneration of the tissue after anti-tumor therapy.

Moreover, manipulation of *hedgehog* signaling properties at different points may be useful as part of a strategy for reshaping/repairing pancreatic tissue both *in vivo* and *in*

vitro. In one embodiment, the present invention makes use of the apparent involvement of *ptc*, *hedgehog*, and *smoothened* in regulating the development of pancreatic tissue. In general, the subject method can be employed therapeutically to regulate the pancreas after physical, chemical or pathological insult. In yet another embodiment, the subject method can be applied to to cell culture techniques, and in particular, may be employed to enhance the initial generation of prosthetic pancreatic tissue devices. Manipulation of proliferation and differentiation of pancreatic tissue, for example, by altering *hedgehog* activity, can provide a means for more carefully controlling the characteristics of a cultured tissue. In an exemplary embodiment, the subject method can be used to augment production of prosthetic devices which require β-islet cells, such as may be used in the encapsulation devices described in, for example, the Aebischer et al. U.S. Patent No. 4,892,538, the Aebischer et al. U.S. Patent No. 5,106,627, the Lim U.S. Patent No. 4,391,909, and the Sefton U.S. Patent No. 4,353,888. Early progenitor cells to the pancreatic islets are multipotential, and apparently coactivate all the islet-specific genes from the time they first appear. As development proceeds, expression of islet-specific hormones, such as insulin, becomes restricted to the pattern of expression characteristic of mature islet cells. The phenotype of mature islet cells, however, is not stable in culture, as reappearence of embryonal traits in mature β-cells can be observed. By utilizing the subject hedgehog antagonists, the differentiation path or proliferative index of the cells can be regulated.

Furthermore, manipulation of the differentiative state of pancreatic tissue can be utilized in conjunction with transplantation of artificial pancreas so as to promote implantation, vascularization, and *in vivo* differentiation and maintenance of the engrafted tissue. For instance, manipulation of *hedgehog* function to affect tissue differentiation can be utilized as a means of maintaining graft viability.

Bellusci et al. (1997) Development 124:53 report that *Sonic hedgehog* regulates lung mesenchymal cell proliferation in vivo. Accordingly, the present method can be used to regulate regeneration of lung tissue, e.g., in the treatment of emphysema.

Fujita et al. (1997) *Biochem Biophys Res Commun* 238:658 reported that Sonic hedgehog is expressed in human lung squamous carcinoma and adenocarcinoma cells. The expression of Sonic hedgehog was also detected in the human lung squamous carcinoma tissues, but not in the normal lung tissue of the same patient. They also observed that Sonic hedgehog stimulates the incorporation of BrdU into the carcinoma cells and stimulates their cell growth, while anti-Shh-N inhibited their cell growth. These results suggest that a *ptc*, *hedgehog*, and/or *smoothened* is involved in the cell growth of such transformed lung tissue and therefore indicates that the subject method can be used as part of a treatment of lung carcinoma and adenocarcinomas, and other proliferative disorders involving the lung epithelia.

Many other tumors may, based on evidence such as involvement of the hedgehog pathway in these tumors, or detected expression of hedgehog or its receptor in these tissues during development, be affected by treatment with the subject compounds. Such tumors include, but are by no means limited to, tumors related to Gorlin's syndrome (e.g., medulloblastoma, meningioma, etc.), tumors evidenced in *pct* knock-out mice (e.g., hemangioma, rhabdomyosarcoma, etc.), tumors resulting from *gli-1* amplification (e.g., glioblastoma, sarcoma, etc.), tumors connected with TRC8, a *ptc* homolog (e.g., renal carcinoma, thyroid carcinoma, etc.), *Ext-1*-related tumors (e.g., bone cancer, etc.), Shh-induced tumors (e.g., lung cancer, chondrosarcomas, etc.), and other tumors (e.g., breast cancer, urogenital cancer (e.g., kidney, bladder, ureter, prostate, etc.), adrenal cancer, gastrointestinal cancer (e.g., stomach, intestine, etc.), etc.).

In still another embodiment of the present invention, compositions comprising *hedgehog* antagonists can be used in the *in vitro* generation of skeletal tissue, such as from skeletogenic stem cells, as well as the *in vivo* treatment of skeletal tissue deficiencies. The present invention particularly contemplates the use of *hedgehog* antagonists to regulate the rate of chondrogenesis and/or osteogenesis. By "skeletal tissue deficiency", it is meant a deficiency in bone or other skeletal connective tissue at any site where it is desired to restore the bone or connective tissue, no matter how the deficiency originated, e.g. whether as a result of surgical intervention, removal of tumor, ulceration, implant, fracture, or other traumatic or degenerative conditions.

For instance, the method of the present invention can be used as part of a regimen for restoring cartilage function to a connective tissue. Such methods are useful in, for example, the repair of defects or lesions in cartilage tissue which is the result of degenerative wear such as that which results in arthritis, as well as other mechanical derangements which may be caused by trauma to the tissue, such as a displacement of torn meniscus tissue, meniscectomy, a laxation of a joint by a torn ligament, malignment of joints, bone fracture, or by hereditary disease. The present reparative method is also useful for remodeling cartilage matrix, such as in plastic or reconstructive surgery, as well as periodontal surgery. The present method may also be applied to improving a previous reparative procedure, for example, following surgical repair of a meniscus, ligament, or cartilage. Furthermore, it may prevent the onset or exacerbation of degenerative disease if applied early enough after trauma.

In one embodiment of the present invention, the subject method comprises treating the afflicted connective tissue with a therapeutically sufficient amount of a *hedgehog* antagonist, particularly an antagonist selective for Indian *hedgehog* signal transduction, to regulate a cartilage repair response in the connective tissue by managing the rate of differentiation and/or proliferation of chondrocytes embedded in the tissue. Such connective tissues as articular cartilage, interarticular cartilage (menisci), costal cartilage (connecting the true ribs and the sternum), ligaments, and tendons are particularly amenable to treatment in reconstructive and/or regenerative therapies using the subject method. As used herein, regenerative therapies include treatment of degenerative states which have progressed to the point of which impairment of the tissue is obviously manifest, as well as preventive treatments of tissue where degeneration is in its earliest stages or imminent.

In an illustrative embodiment, the subject method can be used as part of a therapeutic intervention in the treatment of cartilage of a diarthroidal joint, such as a knee, an ankle, an elbow, a hip, a wrist, a knuckle of either a finger or toe, or a tempomandibular joint. The treatment can be directed to the meniscus of the joint, to the articular cartilage of the joint, or both. To further illustrate, the subject method can be

used to treat a degenerative disorder of a knee, such as which might be the result of traumatic injury (e.g., a sports injury or excessive wear) or osteoarthritis. The subject antagonists may be administered as an injection into the joint with, for instance, an arthroscopic needle. In some instances, the injected agent can be in the form of a hydrogel or other slow release vehicle described above in order to permit a more extended and regular contact of the agent with the treated tissue.

The present invention further contemplates the use of the subject method in the field of cartilage transplantation and prosthetic device therapies. However, problems arise, for instance, because the characteristics of cartilage and fibrocartilage varies between different tissue: such as between articular, meniscal cartilage, ligaments, and tendons, between the two ends of the same ligament or tendon, and between the superficial and deep parts of the tissue. The zonal arrangement of these tissues may reflect a gradual change in mechanical properties, and failure occurs when implanted tissue, which has not differentiated under those conditions, lacks the ability to appropriately respond. For instance, when meniscal cartilage is used to repair anterior cruciate ligaments, the tissue undergoes a metaplasia to pure fibrous tissue. By regulating the rate of chondrogenesis, the subject method can be used to particularly address this problem, by helping to adaptively control the implanted cells in the new environment and effectively resemble hypertrophic chondrocytes of an earlier developmental stage of the tissue.

In similar fashion, the subject method can be applied to enhancing both the generation of prosthetic cartilage devices and to their implantation. The need for improved treatment has motivated research aimed at creating new cartilage that is based on collagen-glycosaminoglycan templates (Stone et al. (1990) *Clin Orthop Relat Red* 252:129), isolated chondrocytes (Grande et al. (1989) *J Orthop Res* 7:208; and Takigawa et al. (1987) *Bone Miner* 2:449), and chondrocytes attached to natural or synthetic polymers (Walitani et al. (1989) *J Bone Jt Surg* 71B:74; Vacanti et al. (1991) *Plast Reconstr Surg* 88:753; von Schroeder et al. (1991) *J Biomed Mater Res* 25:329; Freed et al. (1993) *J Biomed Mater Res* 27:11; and the Vacanti et al. U.S. Patent No. 5,041,138).

For example, chondrocytes can be grown in culture on biodegradable, biocompatible highly porous scaffolds formed from polymers such as polyglycolic acid, polylactic acid, agarose gel, or other polymers which degrade over time as function of hydrolysis of the polymer backbone into innocuous monomers. The matrices are designed to allow adequate nutrient and gas exchange to the cells until engraftment occurs. The cells can be cultured *in vitro* until adequate cell volume and density has developed for the cells to be implanted. One advantage of the matrices is that they can be cast or molded into a desired shape on an individual basis, so that the final product closely resembles the patient's own ear or nose (by way of example), or flexible matrices can be used which allow for manipulation at the time of implantation, as in a joint.

In one embodiment of the subject method, the implants are contacted with a *hedgehog* antagonist during certain stages of the culturing process in order to manage the rate of differentiation of chondrocytes and the formation of hypertrophic chrondrocytes in the culture.

In another embodiment, the implanted device is treated with a *hedgehog* antagonist in order to actively remodel the implanted matrix and to make it more suitable for its intended function. As set out above with respect to tissue transplants, the artificial transplants suffer from the same deficiency of not being derived in a setting which is comparable to the actual mechanical environment in which the matrix is implanted. The ability to regulate the chondrocytes in the matrix by the subject method can allow the implant to acquire characteristics similar to the tissue for which it is intended to replace.

In yet another embodiment, the subject method is used to enhance attachment of prosthetic devices. To illustrate, the subject method can be used in the implantation of a periodontal prosthesis, wherein the treatment of the surrounding connective tissue stimulates formation of periodontal ligament about the prosthesis.

In still further embodiments, the subject method can be employed as part of a regimen for the generation of bone (osteogenesis) at a site in the animal where such skeletal tissue is deficient. Indian hedgehog is particularly associated with the hypertrophic chondrocytes that are ultimately replaced by osteoblasts. For instance,

administration of a *hedgehog* antagonists of the present invention can be employed as part of a method for regulating the rate of bone loss in a subject. For example, preparations comprising *hedgehog* antagonists can be employed, for example, to control endochondral ossification in the formation of a "model" for ossification.

In yet another embodiment of the present invention, a *hedgehog* antagonist can be used to regulate spermatogenesis. The *hedgehog* proteins, particularly Dhh, have been shown to be involved in the differentiation and/or proliferation and maintenance of testicular germ cells. Dhh expression is initiated in Sertoli cell precursors shortly after the activation of Sry (testicular determining gene) and persists in the testis into the adult. Males are viable but infertile, owing to a complete absence of mature sperm. Examination of the developing testis in different genetic backgrounds suggests that Dhh regulates both early and late stages of spermatogenesis. Bitgood et al. (1996) Curr Biol 6:298. In a preferred embodiment, the *hedgehog* antagonist can be used as a contraceptive. In similar fashion, *hedgehog* antagonists of the subject method are potentially useful for modulating normal ovarian function.

The subject method also has wide applicability to the treatment or prophylaxis of disorders afflicting epithelial tissue, as well as in cosmetic uses. In general, the method can be characterized as including a step of administering to an animal an amount of a *hedgehog* antagonist effective to alter the growth state of a treated epithelial tissue. The mode of administration and dosage regimens will vary depending on the epithelial tissue(s) which is to be treated. For example, topical formulations will be preferred where the treated tissue is epidermal tissue, such as dermal or mucosal tissues.

A method which "promotes the healing of a wound" results in the wound healing more quickly as a result of the treatment than a similar wound heals in the absence of the treatment. "Promotion of wound healing" can also mean that the method regulates the proliferation and/or growth of, *inter alia,* keratinocytes, or that the wound heals with less scarring, less wound contraction, less collagen deposition and more superficial surface area. In certain instances, "promotion of wound healing" can also mean that certain

methods of wound healing have improved success rates, (e.g., the take rates of skin grafts,) when used together with the method of the present invention.

Despite significant progress in reconstructive surgical techniques, scarring can be an important obstacle in regaining normal function and appearance of healed skin. This is particularly true when pathologic scarring such as keloids or hypertrophic scars of the hands or face causes functional disability or physical deformity. In the severest circumstances, such scarring may precipitate psychosocial distress and a life of economic deprivation. Wound repair includes the stages of hemostasis, inflammation, proliferation, and remodeling. The proliferative stage involves multiplication of fibroblasts and endothelial and epithelial cells. Through the use of the subject method, the rate of proliferation of epithelial cells in and proximal to the wound can be controlled in order to accelerate closure of the wound and/or minimize the formation of scar tissue.

The present treatment can also be effective as part of a therapeutic regimen for treating oral and paraoral ulcers, e.g. resulting from radiation and/or chemotherapy. Such ulcers commonly develop within days after chemotherapy or radiation therapy. These ulcers usually begin as small, painful irregularly shaped lesions usually covered by a delicate gray necrotic membrane and surrounded by inflammatory tissue. In many instances, lack of treatment results in proliferation of tissue around the periphery of the lesion on an inflammatory basis. For instance, the epithelium bordering the ulcer usually demonstrates proliferative activity, resulting in loss of continuity of surface epithelium. These lesions, because of their size and loss of epithelial integrity, dispose the body to potential secondary infection. Routine ingestion of food and water becomes a very painful event and, if the ulcers proliferate throughout the alimentary canal, diarrhea usually is evident with all its complicating factors. According to the present invention, a treatment for such ulcers which includes application of an *hedgehog* antagonist can reduce the abnormal proliferation and differentiation of the affected epithelium, helping to reduce the severity of subsequent inflammatory events.

The subject method and compositions can also be used to treat wounds resulting from dermatological diseases, such as lesions resulting from autoimmune disorders such

as psoriasis. Atopic dermititis refers to skin trauma resulting from allergies associated with an immune response caused by allergens such as pollens, foods, dander, insect venoms and plant toxins.

In other embodiments, antiproliferative preparations of *hedgehog* antagonists can be used to inhibit lens epithelial cell proliferation to prevent post-operative complications of extracapsular cataract extraction. Cataract is an intractable eye disease and various studies on a treatment of cataract have been made. But at present, the treatment of cataract is attained by surgical operations. Cataract surgery has been applied for a long time and various operative methods have been examined. Extracapsular lens extraction has become the method of choice for removing cataracts. The major medical advantages of this technique over intracapsular extraction are lower incidence of aphakic cystoid macular edema and retinal detachment. Extracapsular extraction is also required for implantation of posterior chamber type intraocular lenses which are now considered to be the lenses of choice in most cases.

However, a disadvantage of extracapsular cataract extraction is the high incidence of posterior lens capsule opacification, often called after-cataract, which can occur in up to 50% of cases within three years after surgery. After-cataract is caused by proliferation of equatorial and anterior capsule lens epithelial cells which remain after extracapsular lens extraction. These cells proliferate to cause Sommerling rings, and along with fibroblasts which also deposit and occur on the posterior capsule, cause opacification of the posterior capsule, which interferes with vision. Prevention of after-cataract would be preferable to treatment. To inhibit secondary cataract formation, the subject method provides a means for inhibiting proliferation of the remaining lens epithelial cells. For example, such cells can be induced to remain quiescent by instilling a solution containing an *hedgehog* antagonist preparation into the anterior chamber of the eye after lens removal. Furthermore, the solution can be osmotically balanced to provide minimal effective dosage when instilled into the anterior chamber of the eye, thereby inhibiting subcapsular epithelial growth with some specificity.

The subject method can also be used in the treatment of corneopathies marked by corneal epithelial cell proliferation, as for example in ocular epithelial disorders such as epithelial downgrowth or squamous cell carcinomas of the ocular surface.

Levine et al. (1997) J Neurosci 17:6277 show that hedgehog proteins can regulate mitogenesis and photoreceptor differentiation in the vertebrate retina, and Ihh is a candidate factor from the pigmented epithelium to promote retinal progenitor proliferation and photoreceptor differentiation. Likewise, Jensen et al. (1997) Development 124:363 demonstrated that treatment of cultures of perinatal mouse retinal cells with the amino-terminal fragment of Sonic hedgehog protein results in an increase in the proportion of cells that incorporate bromodeoxuridine, in total cell numbers, and in rod photoreceptors, amacrine cells and Muller glial cells, suggesting that Sonic hedgehog promotes the proliferation of retinal precursor cells. Thus, the subject method can be used in the treatment of proliferative diseases of retinal cells and regulate photoreceptor differentiation.

Yet another aspect of the present invention relates to the use of the subject method to control hair growth. Hair is basically composed of keratin, a tough and insoluble protein; its chief strength lies in its disulphide bond of cystine. Each individual hair comprises a cylindrical shaft and a root, and is contained in a follicle, a flask-like depression in the skin. The bottom of the follicle contains a finger-like projection termed the papilla, which consists of connective tissue from which hair grows, and through which blood vessels supply the cells with nourishment. The shaft is the part that extends outwards from the skin surface, whilst the root has been described as the buried part of the hair. The base of the root expands into the hair bulb, which rests upon the papilla. Cells from which the hair is produced grow in the bulb of the follicle; they are extruded in the form of fibers as the cells proliferate in the follicle. Hair "growth" refers to the formation and elongation of the hair fiber by the dividing cells.

As is well known in the art, the common hair cycle is divided into three stages: anagen, catagen and telogen. During the active phase (anagen), the epidermal stem cells of the dermal papilla divide rapidly. Daughter cells move upward and differentiate to

form the concentric layers of the hair itself. The transitional stage, catagen, is marked by the cessation of mitosis of the stem cells in the follicle. The resting stage is known as telogen, where the hair is retained within the scalp for several weeks before an emerging new hair developing below it dislodges the telogen-phase shaft from its follicle. From this model it has become clear that the larger the pool of dividing stem cells that differentiate into hair cells, the more hair growth occurs. Accordingly, methods for increasing or reducing hair growth can be carried out by potentiating or inhibiting, respectively, the proliferation of these stem cells.

In certain embodiments, the subject method can be employed as a way of reducing the growth of human hair as opposed to its conventional removal by cutting, shaving, or depilation. For instance, the present method can be used in the treatment of trichosis characterized by abnormally rapid or dense growth of hair, e.g. hypertrichosis. In an exemplary embodiment, *hedgehog* antagonists can be used to manage hirsutism, a disorder marked by abnormal hairiness. The subject method can also provide a process for extending the duration of depilation.

Moreover, because a *hedgehog* antagonist will often be cytostatic to epithelial cells, rather than cytotoxic, such agents can be used to protect hair follicle cells from cytotoxic agents which require progression into S-phase of the cell-cycle for efficacy, e.g. radiation-induced death. Treatment by the subject method can provide protection by causing the hair follicle cells to become quiescent, e.g., by inhibiting the cells from entering S phase, and thereby preventing the follicle cells from undergoing mitotic catastrophe or programmed cell death. For instance, *hedgehog* antagonists can be used for patients undergoing chemo- or radiation-therapies which ordinarily result in hair loss. By inhibiting cell-cycle progression during such therapies, the subject treatment can protect hair follicle cells from death which might otherwise result from activation of cell death programs. After the therapy has concluded, the instant method can also be removed with concommitant relief of the inhibition of follicle cell proliferation.

The subject method can also be used in the treatment of folliculitis, such as folliculitis decalvans, folliculitis ulerythematosa reticulata or keloid folliculitis. For

example, a cosmetic prepration of an *hedgehog* antagonist can be applied topically in the treatment of pseudofolliculitis, a chronic disorder occurring most often in the submandibular region of the neck and associated with shaving, the characteristic lesions of which are erythematous papules and pustules containing buried hairs.

In another aspect of the invention, the subject method can be used to induce differentiation and/or inhibit proliferation of epithelially derived tissue. Such forms of these molecules can provide a basis for differentiation therapy for the treatment of hyperplastic and/or neoplastic conditions involving epithelial tissue. For example, such preparations can be used for the treatment of cutaneous diseases in which there is abnormal proliferation or growth of cells of the skin.

For instance, the pharmaceutical preparations of the invention are intended for the treatment of hyperplastic epidermal conditions, such as keratosis, as well as for the treatment of neoplastic epidermal conditions such as those characterized by a high proliferation rate for various skin cancers, as for example squamous cell carcinoma. The subject method can also be used in the treatment of autoimmune diseases affecting the skin, in particular, of dermatological diseases involving morbid proliferation and/or keratinization of the epidermis, as for example, caused by psoriasis or atopic dermatosis.

Many common diseases of the skin, such as psoriasis, squamous cell carcinoma, keratoacanthoma and actinic keratosis are characterized by localized abnormal proliferation and growth. For example, in psoriasis, which is characterized by scaly, red, elevated plaques on the skin, the keratinocytes are known to proliferate much more rapidly than normal and to differentiate less completely.

In one embodiment, the preparations of the present invention are suitable for the treatment of dermatological ailments linked to keratinization disorders causing abnormal proliferation of skin cells, which disorders may be marked by either inflammatory or non-inflammatory components. To illustrate, therapeutic preparations of a *hedgehog* antagonist, e.g., which promotes quiescense or differentiation can be used to treat varying forms of psoriasis, be they cutaneous, mucosal or ungual. Psoriasis, as described above, is typically characterized by epidermal keratinocytes which display marked proliferative

activation and differentiation along a "regenerative" pathway. Treatment with an antiproliferative embodiment of the subject method can be used to reverse the pathological epidermal activiation and can provide a basis for sustained remission of the disease.

A variety of other keratotic lesions are also candidates for treatment with the subject method. Actinic keratoses, for example, are superficial inflammatory premalignant tumors arising on sun-exposed and irradiated skin. The lesions are erythematous to brown with variable scaling. Current therapies include excisional and cryosurgery. These treatments are painful, however, and often produce cosmetically unacceptable scarring. Accordingly, treatment of keratosis, such as actinic keratosis, can include application, preferably topical, of a *hedgehog* antagonist composition in amounts sufficient to inhibit hyperproliferation of epidermal/epidermoid cells of the lesion.

Acne represents yet another dermatologic ailment which may be treated by the subject method. Acne vulgaris, for instance, is a multifactorial disease most commonly occurring in teenagers and young adults, and is characterized by the appearance of inflammatory and noninflammatory lesions on the face and upper trunk. The basic defect which gives rise to acne vulgaris is hypercornification of the duct of a hyperactive sebaceous gland. Hypercornification blocks the normal mobility of skin and follicle microorganisms, and in so doing, stimulates the release of lipases by *Propinobacterium acnes* and *Staphylococcus epidermidis* bacteria and *Pitrosporum ovale*, a yeast. Treatment with an antiproliferative *hedgehog* antagonist, particularly topical preparations, may be useful for preventing the transitional features of the ducts, e.g. hypercornification, which lead to lesion formation. The subject treatment may further include, for example, antibiotics, retinoids and antiandrogens.

The present invention also provides a method for treating various forms of dermatitis. Dermatitis is a descriptive term referring to poorly demarcated lesions which are either pruritic, erythematous, scaley, blistered, weeping, fissured or crusted. These lesions arise from any of a wide variety of causes. The most common types of dermatitis are atopic, contact and diaper dermatitis. For instance, seborrheic dermatitis is a chronic, usually pruritic, dermatitis with erythema, dry, moist, or greasy scaling, and yellow

crusted patches on various areas, especially the scalp, with exfoliation of an excessive amount of dry scales. The subject method can also be used in the treatment of stasis dermatitis, an often chronic, usually eczematous dermatitis. Actinic dermatitis is dermatitis that due to exposure to actinic radiation such as that from the sun, ultraviolet waves or x- or gamma-radiation. According to the present invention, the subject method can be used in the treatment and/or prevention of certain symptoms of dermatitis caused by unwanted proliferation of epithelial cells. Such therapies for these various forms of dermatitis can also include topical and systemic corticosteroids, antipuritics, and antibiotics.

Ailments which may be treated by the subject method are disorders specific to non-humans, such as mange.

In still another embodiment, the subject method can be used in the treatment of human cancers, such as tumors of epithelial tissues such as the skin. For example, *hedgehog* antagonists can be employed, in the subject method, as part of a treatment for human carcinomas, adenocarcinomas, sarcomas and the like.

In another aspect, the present invention provides pharmaceutical preparations comprising *hedgehog* antagonists. The *hedgehog* antagonists for use in the subject method may be conveniently formulated for administration with a biologically acceptable medium, such as water, buffered saline, polyol (for example, glycerol, propylene glycol, liquid polyethylene glycol and the like) or suitable mixtures thereof. The optimum concentration of the active ingredient(s) in the chosen medium can be determined empirically, according to procedures well known to medicinal chemists. As used herein, "biologically acceptable medium" includes any and all solvents, dispersion media, and the like which may be appropriate for the desired route of administration of the pharmaceutical preparation. The use of such media for pharmaceutically active substances is known in the art. Except insofar as any conventional media or agent is incompatible with the activity of the *hedgehog* antagonist, its use in the pharmaceutical preparation of the invention is contemplated. Suitable vehicles and their formulation inclusive of other proteins are described, for example, in the book *Remington's*

Pharmaceutical Sciences (Remington's Pharmaceutical Sciences. Mack Publishing Company, Easton, Pa., USA 1985). These vehicles include injectable "deposit formulations".

Pharmaceutical formulations of the present invention can also include veterinary compositions, e.g., pharmaceutical preparations of the *hedgehog* antagonists suitable for veterinary uses, e.g., for the treatment of live stock or domestic animals, e.g., dogs.

Methods of introduction may also be provided by rechargeable or biodegradable devices. Various slow release polymeric devices have been developed and tested *in vivo* in recent years for the controlled delivery of drugs, including proteinacious biopharmaceuticals. A variety of biocompatible polymers (including hydrogels), including both biodegradable and non-degradable polymers, can be used to form an implant for the sustained release of a *hedgehog* antagonist at a particular target site.

The preparations of the present invention may be given orally, parenterally, topically, or rectally. They are of course given by forms suitable for each administration route. For example, they are administered in tablets or capsule form, by injection, inhalation, eye lotion, ointment, suppository, controlled release patch, etc. administration by injection, infusion or inhalation; topical by lotion or ointment; and rectal by suppositories. Oral and topical administrations are preferred.

The phrases "parenteral administration" and "administered parenterally" as used herein means modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticulare, subcapsular, subarachnoid, intraspinal and intrasternal injection and infusion.

The phrases "systemic administration," "administered systemically," "peripheral administration" and "administered peripherally" as used herein mean the administration of a compound, drug or other material other than directly into the central nervous system, such that it enters the patient's system and, thus, is subject to metabolism and other like processes, for example, subcutaneous administration.

These compounds may be administered to humans and other animals for therapy by any suitable route of administration, including orally, nasally, as by, for example, a spray, rectally, intravaginally, parenterally, intracisternally and topically, as by powders, ointments or drops, including buccally and sublingually.

Regardless of the route of administration selected, the compounds of the present invention, which may be used in a suitable hydrated form, and/or the pharmaceutical compositions of the present invention, are formulated into pharmaceutically acceptable dosage forms such as described below or by other conventional methods known to those of skill in the art.

Actual dosage levels of the active ingredients in the pharmaceutical compositions of this invention may be varied so as to obtain an amount of the active ingredient which is effective to achieve the desired therapeutic response for a particular patient, composition, and mode of administration, without being toxic to the patient.

The selected dosage level will depend upon a variety of factors including the activity of the particular compound of the present invention employed, or the ester, salt or amide thereof, the route of administration, the time of administration, the rate of excretion of the particular compound being employed, the duration of the treatment, other drugs, compounds and/or materials used in combination with the particular hedgehog antagonist employed, the age, sex, weight, condition, general health and prior medical history of the patient being treated, and like factors well known in the medical arts.

A physician or veterinarian having ordinary skill in the art can readily determine and prescribe the effective amount of the pharmaceutical composition required. For example, the physician or veterinarian could start doses of the compounds of the invention employed in the pharmaceutical composition at levels lower than that required in order to achieve the desired therapeutic effect and gradually increase the dosage until the desired effect is achieved.

In general, a suitable daily dose of a compound of the invention will be that amount of the compound which is the lowest dose effective to produce a therapeutic effect. Such an effective dose will generally depend upon the factors described above.

Generally, intravenous, intracerebroventricular and subcutaneous doses of the compounds of this invention for a patient will range from about 0.0001 to about 100 mg per kilogram of body weight per day.

If desired, the effective daily dose of the active compound may be administered as two, three, four, five, six or more sub-doses administered separately at appropriate intervals throughout the day, optionally, in unit dosage forms.

The term "treatment" is intended to encompass also prophylaxis, therapy and cure.

The patient receiving this treatment is any animal in need, including primates, in particular humans, and other mammals such as equines, cattle, swine and sheep; and poultry and pets in general.

The compound of the invention can be administered as such or in admixtures with pharmaceutically acceptable and/or sterile carriers and can also be administered in conjunction with other antimicrobial agents such as penicillins, cephalosporins, aminoglycosides and glycopeptides. Conjunctive therapy, thus includes sequential, simultaneous and separate administration of the active compound in a way that the therapeutical effects of the first administered one is not entirely disappeared when the subsequent is administered.

V. Pharmaceutical Compositions

While it is possible for a compound of the present invention to be administered alone, it is preferable to administer the compound as a pharmaceutical formulation (composition). The *hedgehog* antagonists according to the invention may be formulated for administration in any convenient way for use in human or veterinary medicine. In certain embodiments, the compound included in the pharmaceutical preparation may be active itself, or may be a prodrug, e.g., capable of being converted to an active compound in a physiological setting.

Thus, another aspect of the present invention provides pharmaceutically acceptable compositions comprising a therapeutically effective amount of one or more of

the compounds described above, formulated together with one or more pharmaceutically acceptable carriers (additives) and/or diluents. As described in detail below, the pharmaceutical compositions of the present invention may be specially formulated for administration in solid or liquid form, including those adapted for the following: (1) oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets, boluses, powders, granules, pastes for application to the tongue; (2) parenteral administration, for example, by subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution or suspension; (3) topical application, for example, as a cream, ointment or spray applied to the skin; or (4) intravaginally or intrarectally, for example, as a pessary, cream or foam. However, in certain embodiments the subject compounds may be simply dissolved or suspended in sterile water. In certain embodiments, the pharmaceutical preparation is non-pyrogenic, i.e., does not elevate the body temperature of a patient.

The phrase "therapeutically effective amount" as used herein means that amount of a compound, material, or composition comprising a compound of the present invention which is effective for producing some desired therapeutic effect by overcoming a *hedgehog* gain-of-function phenotype in at least a sub-population of cells in an animal and thereby blocking the biological consequences of that pathway in the treated cells, at a reasonable benefit/risk ratio applicable to any medical treatment.

The phrase "pharmaceutically acceptable" is employed herein to refer to those compounds, materials, compositions, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

The phrase "pharmaceutically acceptable carrier" as used herein means a pharmaceutically acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, solvent or encapsulating material, involved in carrying or transporting the subject antagonists from one organ, or portion of the body, to another organ, or portion of the body. Each carrier must be "acceptable" in the sense of being

compatible with the other ingredients of the formulation and not injurious to the patient. Some examples of materials which can serve as pharmaceutically acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) phosphate buffer solutions; and (21) other non-toxic compatible substances employed in pharmaceutical formulations.

As set out above, certain embodiments of the present *hedgehog* antagonists may contain a basic functional group, such as amino or alkylamino, and are, thus, capable of forming pharmaceutically acceptable salts with pharmaceutically acceptable acids. The term "pharmaceutically acceptable salts" in this respect, refers to the relatively non-toxic, inorganic and organic acid addition salts of compounds of the present invention. These salts can be prepared *in situ* during the final isolation and purification of the compounds of the invention, or by separately reacting a purified compound of the invention in its free base form with a suitable organic or inorganic acid, and isolating the salt thus formed. Representative salts include the hydrobromide, hydrochloride, sulfate, bisulfate, phosphate, nitrate, acetate, valerate, oleate, palmitate, stearate, laurate, benzoate, lactate, phosphate, tosylate, citrate, maleate, fumarate, succinate, tartrate, napthylate, mesylate, glucoheptonate, lactobionate, and laurylsulphonate salts and the like. (See, for example, Berge et al. (1977) "Pharmaceutical Salts", *J. Pharm. Sci.* 66:1-19)

The pharmaceutically acceptable salts of the subject compounds include the conventional nontoxic salts or quaternary ammonium salts of the compounds, e.g., from non-toxic organic or inorganic acids. For example, such conventional nontoxic salts include those derived from inorganic acids such as hydrochloride, hydrobromic, sulfuric,

sulfamic, phosphoric, nitric, and the like; and the salts prepared from organic acids such as acetic, propionic, succinic, glycolic, stearic, lactic, malic, tartaric, citric, ascorbic, palmitic, maleic, hydroxymaleic, phenylacetic, glutamic, benzoic, salicyclic, sulfanilic, 2-acetoxybenzoic, fumaric, toluenesulfonic, methanesulfonic, ethane disulfonic, oxalic, isothionic, and the like.

In other cases, the compounds of the present invention may contain one or more acidic functional groups and, thus, are capable of forming pharmaceutically acceptable salts with pharmaceutically acceptable bases. The term "pharmaceutically acceptable salts" in these instances refers to the relatively non-toxic, inorganic and organic base addition salts of compounds of the present invention. These salts can likewise be prepared *in situ* during the final isolation and purification of the compounds, or by separately reacting the purified compound in its free acid form with a suitable base, such as the hydroxide, carbonate or bicarbonate of a pharmaceutically acceptable metal cation, with ammonia, or with a pharmaceutically acceptable organic primary, secondary or tertiary amine. Representative alkali or alkaline earth salts include the lithium, sodium, potassium, calcium, magnesium, and aluminum salts and the like. Representative organic amines useful for the formation of base addition salts include ethylamine, diethylamine, ethylenediamine, ethanolamine, diethanolamine, piperazine and the like. (See, for example, Berge et al., *supra*)

Wetting agents, emulsifiers and lubricants, such as sodium lauryl sulfate and magnesium stearate, as well as coloring agents, release agents, coating agents, sweetening, flavoring and perfuming agents, preservatives and antioxidants can also be present in the compositions.

Examples of pharmaceutically acceptable antioxidants include: (1) water soluble antioxidants, such as ascorbic acid, cysteine hydrochloride, sodium bisulfate, sodium metabisulfite, sodium sulfite and the like; (2) oil-soluble antioxidants, such as ascorbyl palmitate, butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), lecithin, propyl gallate, alpha-tocopherol, and the like; and (3) metal chelating agents, such as

citric acid, ethylenediamine tetraacetic acid (EDTA), sorbitol, tartaric acid, phosphoric acid, and the like.

Formulations of the present invention include those suitable for oral, nasal, topical (including buccal and sublingual), rectal, vaginal and/or parenteral administration. The formulations may conveniently be presented in unit dosage form and may be prepared by any methods well known in the art of pharmacy. The amount of active ingredient which can be combined with a carrier material to produce a single dosage form will vary depending upon the host being treated, the particular mode of administration. The amount of active ingredient which can be combined with a carrier material to produce a single dosage form will generally be that amount of the compound which produces a therapeutic effect. Generally, out of one hundred per cent, this amount will range from about 1 per cent to about ninety-nine percent of active ingredient, preferably from about 5 per cent to about 70 per cent, most preferably from about 10 per cent to about 30 per cent.

Methods of preparing these formulations or compositions include the step of bringing into association a compound of the present invention with the carrier and, optionally, one or more accessory ingredients. In general, the formulations are prepared by uniformly and intimately bringing into association a compound of the present invention with liquid carriers, or finely divided solid carriers, or both, and then, if necessary, shaping the product.

Formulations of the invention suitable for oral administration may be in the form of capsules, cachets, pills, tablets, lozenges (using a flavored basis, usually sucrose and acacia or tragacanth), powders, granules, or as a solution or a suspension in an aqueous or non-aqueous liquid, or as an oil-in-water or water-in-oil liquid emulsion, or as an elixir or syrup, or as pastilles (using an inert base, such as gelatin and glycerin, or sucrose and acacia) and/or as mouth washes and the like, each containing a predetermined amount of a compound of the present invention as an active ingredient. A compound of the present invention may also be administered as a bolus, electuary or paste.

In solid dosage forms of the invention for oral administration (capsules, tablets, pills, dragees, powders, granules and the like), the active ingredient is mixed with one or

more pharmaceutically acceptable carriers, such as sodium citrate or dicalcium phosphate, and/or any of the following: (1) fillers or extenders, such as starches, lactose, sucrose, glucose, mannitol, and/or silicic acid; (2) binders, such as, for example, carboxymethylcellulose, alginates, gelatin, polyvinyl pyrrolidone, sucrose and/or acacia; (3) humectants, such as glycerol; (4) disintegrating agents, such as agar-agar, calcium carbonate, potato or tapioca starch, alginic acid, certain silicates, and sodium carbonate; (5) solution retarding agents, such as paraffin; (6) absorption accelerators, such as quaternary ammonium compounds; (7) wetting agents, such as, for example, cetyl alcohol and glycerol monostearate; (8) absorbents, such as kaolin and bentonite clay; (9) lubricants, such a talc, calcium stearate, magnesium stearate, solid polyethylene glycols, sodium lauryl sulfate, and mixtures thereof; and (10) coloring agents. In the case of capsules, tablets and pills, the pharmaceutical compositions may also comprise buffering agents. Solid compositions of a similar type may also be employed as fillers in soft and hard-filled gelatin capsules using such excipients as lactose or milk sugars, as well as high molecular weight polyethylene glycols and the like.

A tablet may be made by compression or molding, optionally with one or more accessory ingredients. Compressed tablets may be prepared using binder (for example, gelatin or hydroxypropylmethyl cellulose), lubricant, inert diluent, preservative, disintegrant (for example, sodium starch glycolate or cross-linked sodium carboxymethyl cellulose), surface-active or dispersing agent. Molded tablets may be made by molding in a suitable machine a mixture of the powdered compound moistened with an inert liquid diluent.

The tablets, and other solid dosage forms of the pharmaceutical compositions of the present invention, such as dragees, capsules, pills and granules, may optionally be scored or prepared with coatings and shells, such as enteric coatings and other coatings well known in the pharmaceutical-formulating art. They may also be formulated so as to provide slow or controlled release of the active ingredient therein using, for example, hydroxypropylmethyl cellulose in varying proportions to provide the desired release profile, other polymer matrices, liposomes and/or microspheres. They may be sterilized by, for example, filtration through a bacteria-retaining filter, or by incorporating

sterilizing agents in the form of sterile solid compositions which can be dissolved in sterile water, or some other sterile injectable medium immediately before use. These compositions may also optionally contain opacifying agents and may be of a composition that they release the active ingredient(s) only, or preferentially, in a certain portion of the gastrointestinal tract, optionally, in a delayed manner. Examples of embedding compositions which can be used include polymeric substances and waxes. The active ingredient can also be in micro-encapsulated form, if appropriate, with one or more of the above-described excipients.

Liquid dosage forms for oral administration of the compounds of the invention include pharmaceutically acceptable emulsions, microemulsions, solutions, suspensions, syrups and elixirs. In addition to the active ingredient, the liquid dosage forms may contain inert diluents commonly used in the art, such as, for example, water or other solvents, solubilizing agents and emulsifiers, such as ethyl alcohol, isopropyl alcohol, ethyl carbonate, ethyl acetate, benzyl alcohol, benzyl benzoate, propylene glycol, 1,3-butylene glycol, oils (in particular, cottonseed, groundnut, corn, germ, olive, castor and sesame oils), glycerol, tetrahydrofuryl alcohol, polyethylene glycols and fatty acid esters of sorbitan, and mixtures thereof.

Besides inert diluents, the oral compositions can also include adjuvants such as wetting agents, emulsifying and suspending agents, sweetening, flavoring, coloring, perfuming and preservative agents.

Suspensions, in addition to the active compounds, may contain suspending agents as, for example, ethoxylated isostearyl alcohols, polyoxyethylene sorbitol and sorbitan esters, microcrystalline cellulose, aluminum metahydroxide, bentonite, agar-agar and tragacanth, and mixtures thereof.

It is known that sterols, such as cholesterol, will form complexes with cyclodextrins. Thus, in preferred embodiments, where the inhibitor is a steroidal alkaloid, it may be formulated with cyclodextrins, such as α-, β- and γ-cyclodextrin, dimethyl-β cyclodextrin and 2-hydroxypropyl-β-cyclodextrin.

Formulations of the pharmaceutical compositions of the invention for rectal or vaginal administration may be presented as a suppository, which may be prepared by mixing one or more compounds of the invention with one or more suitable nonirritating excipients or carriers comprising, for example, cocoa butter, polyethylene glycol, a suppository wax or a salicylate, and which is solid at room temperature, but liquid at body temperature and, therefore, will melt in the rectum or vaginal cavity and release the active *hedgehog* antagonist.

Formulations of the present invention which are suitable for vaginal administration also include pessaries, tampons, creams, gels, pastes, foams or spray formulations containing such carriers as are known in the art to be appropriate.

Dosage forms for the topical or transdermal administration of a compound of this invention include powders, sprays, ointments, pastes, creams, lotions, gels, solutions, patches and inhalants. The active compound may be mixed under sterile conditions with a pharmaceutically acceptable carrier, and with any preservatives, buffers, or propellants which may be required.

The ointments, pastes, creams and gels may contain, in addition to an active compound of this invention, excipients, such as animal and vegetable fats, oils, waxes, paraffins, starch, tragacanth, cellulose derivatives, polyethylene glycols, silicones, bentonites, silicic acid, talc and zinc oxide, or mixtures thereof.

Powders and sprays can contain, in addition to a compound of this invention, excipients such as lactose, talc, silicic acid, aluminum hydroxide, calcium silicates and polyamide powder, or mixtures of these substances. Sprays can additionally contain customary propellants, such as chlorofluorohydrocarbons and volatile unsubstituted hydrocarbons, such as butane and propane.

Transdermal patches have the added advantage of providing controlled delivery of a compound of the present invention to the body. Such dosage forms can be made by dissolving or dispersing the *hedgehog* antagonists in the proper medium. Absorption enhancers can also be used to increase the flux of the *hedgehog* antagonists across the

skin. The rate of such flux can be controlled by either providing a rate controlling membrane or dispersing the compound in a polymer matrix or gel.

Ophthalmic formulations, eye ointments, powders, solutions and the like, are also contemplated as being within the scope of this invention.

Pharmaceutical compositions of this invention suitable for parenteral administration comprise one or more compounds of the invention in combination with one or more pharmaceutically acceptable sterile isotonic aqueous or nonaqueous solutions, dispersions, suspensions or emulsions, or sterile powders which may be reconstituted into sterile injectable solutions or dispersions just prior to use, which may contain antioxidants, buffers, bacteriostats, solutes which render the formulation isotonic with the blood of the intended recipient or suspending or thickening agents.

Examples of suitable aqueous and nonaqueous carriers which may be employed in the pharmaceutical compositions of the invention include water, ethanol, polyols (such as glycerol, propylene glycol, polyethylene glycol, and the like), and suitable mixtures thereof, vegetable oils, such as olive oil, and injectable organic esters, such as ethyl oleate. Proper fluidity can be maintained, for example, by the use of coating materials, such as lecithin, by the maintenance of the required particle size in the case of dispersions, and by the use of surfactants.

These compositions may also contain adjuvants such as preservatives, wetting agents, emulsifying agents and dispersing agents. Prevention of the action of microorganisms may be ensured by the inclusion of various antibacterial and antifungal agents, for example, paraben, chlorobutanol, phenol sorbic acid, and the like. It may also be desirable to include isotonic agents, such as sugars, sodium chloride, and the like into the compositions. In addition, prolonged absorption of the injectable pharmaceutical form may be brought about by the inclusion of agents which delay absorption such as aluminum monostearate and gelatin.

In some cases, in order to prolong the effect of a drug, it is desirable to slow the absorption of the drug from subcutaneous or intramuscular injection. This may be accomplished by the use of a liquid suspension of crystalline or amorphous material

having poor water solubility. The rate of absorption of the drug then depends upon its rate of dissolution which, in turn, may depend upon crystal size and crystalline form. Alternatively, delayed absorption of a parenterally administered drug form is accomplished by dissolving or suspending the drug in an oil vehicle.

Injectable depot forms are made by forming microencapsule matrices of the subject compounds in biodegradable polymers such as polylactide-polyglycolide. Depending on the ratio of drug to polymer, and the nature of the particular polymer employed, the rate of drug release can be controlled. Examples of other biodegradable polymers include poly(orthoesters) and poly(anhydrides). Depot injectable formulations are also prepared by entrapping the drug in liposomes or microemulsions which are compatible with body tissue.

When the compounds of the present invention are administered as pharmaceuticals, to humans and animals, they can be given per se or as a pharmaceutical composition containing, for example, 0.1 to 99.5% (more preferably, 0.5 to 90%) of active ingredient in combination with a pharmaceutically acceptable carrier.

The addition of the active compound of the invention to animal feed is preferably accomplished by preparing an appropriate feed premix containing the active compound in an effective amount and incorporating the premix into the complete ration.

Alternatively, an intermediate concentrate or feed supplement containing the active ingredient can be blended into the feed. The way in which such feed premixes and complete rations can be prepared and administered are described in reference books (such as "Applied Animal Nutrition", W.H. Freedman and CO., San Francisco, U.S.A., 1969 or "Livestock Feeds and Feeding" O and B books, Corvallis, Ore., U.S.A., 1977).

VI. Synthetic Schemes and Identification of Active Antagonists

The subjects steroidal alkaloids, and congeners thereof, can be prepared readily by employing the cross-coupling technologies of Suzuki, Stille, and the like. These coupling

reactions are carried out under relatively mild conditions and tolerate a wide range of "spectator" functionality.

a. Combinatorial Libraries

The compounds of the present invention, particularly libraries of variants having various representative classes of substituents, are amenable to combinatorial chemistry and other parallel synthesis schemes (see, for example, PCT WO 94/08051). The result is that large libraries of related compounds, e.g. a variegated library of compounds represented above, can be screened rapidly in high throughput assays in order to identify potential *hedgehog* antagonist lead compounds, as well as to refine the specificity, toxicity, and/or cytotoxic-kinetic profile of a lead compound. For instance, *ptc*, *hedgehog*, or *smoothened* bioactivity assays, such as may be developed using cells with either a *ptc* loss-of-function, *hedgehog* gain-of-function, or *smoothened* gain-of-function, can be used to screen a library of the subject compounds for those having agonist activity toward *ptc* or antagonist activity towards *hedgehog* or *smoothened*.

Simply for illustration, a combinatorial library for the purposes of the present invention is a mixture of chemically related compounds which may be screened together for a desired property. The preparation of many related compounds in a single reaction greatly reduces and simplifies the number of screening processes which need to be carried out. Screening for the appropriate physical properties can be done by conventional methods.

Diversity in the library can be created at a variety of different levels. For instance, the substrate aryl groups used in the combinatorial reactions can be diverse in terms of the core aryl moiety, e.g., a variegation in terms of the ring structure, and/or can be varied with respect to the other substituents.

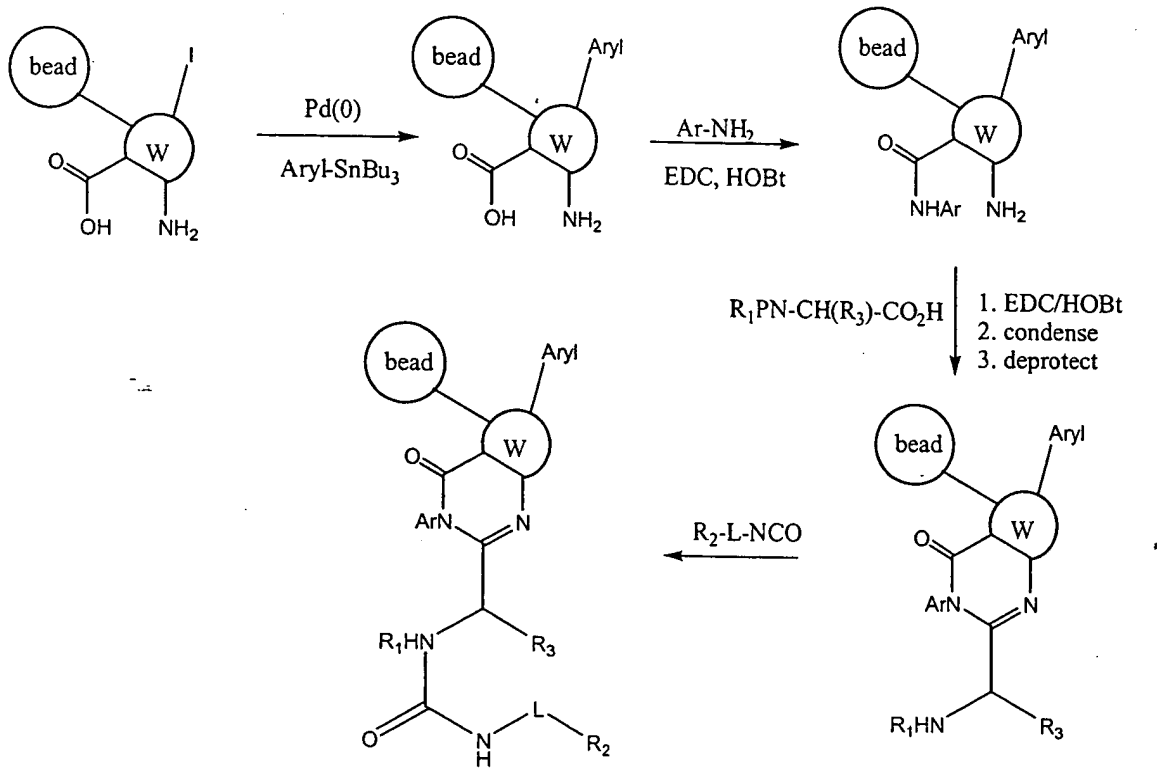
A variety of techniques are available in the art for generating combinatorial libraries of small organic molecules such as the subject *hedgehog* antagonists. See, for example, Blondelle et al. (1995) Trends Anal. Chem. 14:83; the Affymax U.S. Patents 5,359,115 and 5,362,899: the Ellman U.S. Patent 5,288,514: the Still et al. PCT

publication WO 94/08051; the ArQule U.S. Patents 5,736,412 and 5,712,171; Chen et al. (1994) JACS 116:2661: Kerr et al. (1993) JACS 115:252; PCT publications WO92/10092, WO93/09668 and WO91/07087; and the Lerner et al. PCT publication WO93/20242). Accordingly, a variety of libraries on the order of about 100 to 1,000,000 or more diversomers of the subject *hedgehog* antagonists can be synthesized and screened for particular activity or property.

In an exemplary embodiment, a library of candidate *hedgehog* antagonists diversomers can be synthesized utilizing a scheme adapted to the techniques described in the Still et al. PCT publication WO 94/08051, e.g., being linked to a polymer bead by a hydrolyzable or photolyzable group, optionally located at one of the positions of the candidate antagonists or a substituent of a synthetic intermediate. According to the Still et al. technique, the library is synthesized on a set of beads, each bead including a set of tags identifying the particular diversomer on that bead. The bead library can then be "plated" with *ptc* loss-of-function, *hedgehog* gain-of-function, or *smoothened* gain-of-function cells for which an *hedgehog* antagonist is sought. The diversomers can be released from the bead, e.g. by hydrolysis.

The structures of the compounds useful in the present invention lend themselves readily to efficient synthesis. The nature of the structures, as generally described by formulas I and II, allows the assembly of such compounds using some combination of X, Y, and Z moieties, as set forth above. X and Z moieties are heteroatomic in nature, and thus allow the formation of diverse bonds to carbon. Y moieties, when present, take the form of electrophilic moieties, such as carbonyl and sulfonyl groups, and reagents are readily available for the attachment of such moieties to heteroatomic groups like X and Z. The vast majority of such reactions, including those depicted in Figures 11, 12, 15, and 16 are both extremely mild and extremely reliable, and are thus perfectly suited for combinatorial chemistry. The facile nature of such a combinatorial approach towards the generation of a library of test compounds is apparent in the exemplary scheme below (P = protecting group), wherein the various groups of a compound according to Formula II are linked combinatorially (e.g., using one of the methods described above), with combinatorial functionalization of the core ring system (e.g., W) bestowing additional

diversity to the library. Even greater diversity may be attained by, for example, utilizing a range of electrophilic groups when appending a subunit, i.e., using a range of PO-Ar-L-C(O)Cl, PO-Ar-L-NCO, PO-Ar-L-SO_2Cl, etc. when appending the R_2 subunit.



Many variations on the above and related pathways permit the synthesis of widely diverse libraries of compounds which may be tested as inhibitors of *hedgehog* function.

b. Screening Assays

There are a variety of assays available for determining the ability of a compound to agonize *ptc* function or antagonize *smoothened* or *hedgehog* function, many of which can be disposed in high-throughput formats. In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Thus, libraries of synthetic and natural products can be sampled for other compounds which are *hedgehog* antagonists.

In addition to cell-free assays, test compounds can also be tested in cell-based assays. In one embodiment, cell which have a *ptc* loss-of-function, *hedgehog* gain-of-function, or *smoothened* gain-of-function phenotype can be contacted with a test agent of interest, with the assay scoring for, e.g., inhibition of proliferation of the cell in the presence of the test agent.

A number of gene products have been implicated in *patched*-mediated signal transduction, including *patched*, transcription factors of the *cubitus interruptus* (ci) family, the serine/threonine kinase *fused* (fu) and the gene products of *costal-2*, *smoothened* and *suppressor of fused*.

The induction of cells by hedgehog proteins sets in motion a cascade involving the activation and inhibition of downstream effectors, the ultimate consequence of which is, in some instances, a detectable change in the transcription or translation of a gene. Potential transcriptional targets of *hedgehog*-mediated signaling are the *patched* gene (Hidalgo and Ingham, 1990 Development 110, 291-301; Marigo et al., 1996) and the vertebrate homologs of the drosophila cubitus interruptus gene, the *GLI* genes (Hui et al. (1994) Dev Biol 162:402-413). *Patched* gene expression has been shown to be induced in cells of the limb bud and the neural plate that are responsive to *Shh*. (Marigo et al. (1996) PNAS 93:9346-51; Marigo et al. (1996) Development 122:1225-1233). The *Gli* genes encode putative transcription factors having zinc finger DNA binding domains (Orenic et al. (1990) Genes & Dev 4:1053-1067; Kinzler et al. (1990) Mol Cell Biol 10:634-642). Transcription of the *Gli* gene has been reported to be upregulated in response to *hedgehog* in limb buds, while transcription of the *Gli3* gene is downregulated in response to *hedgehog* induction (Marigo et al. (1996) Development 122:1225-1233). By selecting transcriptional regulatory sequences from such target genes, e.g., from *patched* or *Gli* genes, that are responsible for the up- or down-regulation of these genes in response to *hedgehog* signalling, and operatively linking such promoters to a reporter gene, one can derive a transcription based assay which is sensitive to the ability of a specific test compound to modify *hedgehog*-mediated signalling pathways. Expression of the reporter gene, thus, provides a valuable screening tool for the development of compounds that act as antagonists of *hedgehog*.

Reporter gene based assays of this invention measure the end stage of the above described cascade of events, e.g., transcriptional modulation. Accordingly, in practicing one embodiment of the assay, a reporter gene construct is inserted into the reagent cell in order to generate a detection signal dependent on *ptc* loss-of-function, *hedgehog* gain-of-function, *smoothened* gain-of-function, or stimulation by SHH itself. The amount of transcription from the reporter gene may be measured using any method known to those of skill in the art to be suitable. For example, mRNA expression from the reporter gene may be detected using RNAse protection or RNA-based PCR, or the protein product of the reporter gene may be identified by a characteristic stain or an intrinsic biological activity. The amount of expression from the reporter gene is then compared to the amount of expression in either the same cell in the absence of the test compound or it may be compared with the amount of transcription in a substantially identical cell that lacks the target receptor protein. Any statistically or otherwise significant decrease in the amount of transcription indicates that the test compound has in some manner agonized the normal *ptc* signal (or antagonized the gain-of-function *hedgehog* or *smoothened* signal), e.g., the test compound is a potential *hedgehog* antagonist.

Exemplification

The invention now being generally described, it will be more readily understood by reference to the following examples which are included merely for purposes of illustration of certain aspects and embodiments of the present invention, and are not intended to limit the invention.

Lead Compound Discovery/ High-throughput Screening Assay

Compounds to be tested are dissolved in DMSO to a concentration of 10 mM, and stored at -20 °C. To activate the *Hedgehog* pathway in the assay cells, an octylated (lipid-modified) form of the N-terminal fragment of the *Sonic Hedgehog* protein (OCT-SHH) is used. This N-terminal SHH fragment is produced bacterially.

Compounds may be tested in the "Gli-Luc" assay below, using the cell line 10T(s12), wherein the cells contain a Hedgehog-responsive reporter construct utilizing Luciferase as the reporter gene. In this way, *Hedgehog* pathway signaling activity can be measured via the Gli-Luc response.

10t1/2(s12) cells are plated in a 96-well micro-titer plate (MTP) at 20,000 cells/well in full medium [DMEM with 10% FBS]. Then plates are placed in the incubator for incubation overnight (O/N), at 37 °C and 5% CO_2. After 24 h, the medium is replaced with Luciferase-assay medium (DMEM with 0.5% FBS). Compounds are thawed and diluted in assay medium at 3:1000 (about 300-fold) resulting in a starting concentration of about 30 μM.

Subsequently, 150 μl of each 30 μM sample is added to the first wells (in triplicate). The MTP samples are then diluted at 3-fold dilutions to a total of seven wells, ultimately resulting in a regiment of seven dilutions in triplicate, for each compound. Next, the protein ligand OCT-SHH is diluted in Luciferase-assay medium and added to each well at a final concentration of 0.3 μg/ml. Plates are then returned to the incubator for further incubation O/N, at 37 °C and 5% CO_2. After about 24 h, plates are removed from the incubator and the medium is aspirated/discarded. Wells are washed once with assay buffer [PBS + 1 mM Mg^{2+} and 1 mM Ca^{2+}]. Then 50 μl of assay buffer is added to each well. The Luciferase assay reagent is prepared as described by the vendor (LucLite kit from Packard), and 50 μl is added to each well. Plates are incubated at room temperature (RT) for about 30 minutes after which the signals are read, again at RT, on a Topcount (Packard).

Compounds identified in this assay are depicted in Figure 32. The discovery of these Shh-induced Gli-transcription activity inhibitors exemplifies the utility of the claims in this patent. Activities for these compounds are presented in Table 1 below.

Table 1

Compound	IC_{50}	Compound	IC_{50}
31	<10 μM	55	<5 μM
32	<5 μM	56	<10 μM
34	<5 μM	57	<10 μM
11	<5 μM	58	<5 μM
36	<5 μM	59	<5 μM
38	<5 μM	60	<5 μM
39	<5 μM	61	<1 μM
40	<10 μM	62	<1 μM
41	<10 μM	63	<10 μM
42	<5 μM	64	<10 μM
43	<10 μM	65	<10 μM
44	<1 μM	66	<10 μM
45	<5 μM	67	<5 μM
46	<0.5 μM	68	<1 μM
47	<5 μM	69	<0.5 μM
48	<0.5 μM	5	<0.1 μM
49	<1 μM	71	<10 μM
50	<1 μM	6	<0.5 μM
51	<5 μM	73	<5 μM
52	<1 μM	74	<5 μM
53	<1 μM	75	<5 μM
54	<5 μM		

Additional data is presented in Figure 33. Mouse 456 is a Ptc-knockout heterozygote that received UV irradiation for 6 months. The mouse developed many small BCC lesions, which were blue after X-gal staining. The mouse was sacrificed and the skin was excised with a 2 mm skin punch. Those skin punches were then cultured for 6 days. Comparing to vehicle (DMSO), compound **11** can decrease the number and size of BCC lesions (blue spots in the picture). The effect was more obvious when 10 μM of **11** was added to the culture medium instead of 5 μM of **11**. In short, this experiment suggests that **11** is able to inhibit murine BCC lesions in mouse #456.

In yet another experiment, E12.5 old *ptc-1 (d11) lacZ* lungs were harvested and transgenic embryos identified by lacZ detection using tails. Lung explants were grown submerged in mouse explant medium (DMEM based, additives optimized for the culture

of mouse lungs) for 48 hrs, fixed in lacZ fixative, rinsed and stained for lacZ O/N at 37 °C. Control tissue was untreated, while test tissue was treated with **11**. Results are depicted in Figure 34: (A) Untreated control. Strong *lacZ* expression can be observed in distal and proximal mesenchyme. (B) Treatment with **11** leads to significantly decreased reporter gene expression, as evidenced especially by the weak signal surrounding the distal branching tips of the growing lung epithelium.

Preparation of compounds of the present invention

a. Illustrative synthetic schemes

Exemplary synthesis schemes for generating *hedgehog* antagonists useful in the methods and compositions of the present invention are shown in Figures 1-31.

The reaction conditions in the illustrated schemes of Figure 1-31 are as follows:

1) R_1CH_2CN, $NaNH_2$, toluene

(Arzneim-Forsch, 1990, 40, 11, 1242)

2) H_2SO_4, H_2O, reflux

(Arzneim-Forsch, 1990, 40, 11, 1242)

3) H_2SO_4, EtOH, reflux

(Arzneim-Forsch, 1990, 40, 11, 1242)

4) NaOH, EtOH, reflux

5) $(Boc)_2O$, 2M NaOH, THF

6) LiHDMS, R_1X, THF

(Merck Patent Applic # WO 96/06609)

7) Pd-C, H_2, MeOH

8) t-BuONO, CuBr, HBr, H_2O

(J. Org. Chem. 1977, 42, 2426)

9) $ArB(OH)_2$, $Pd(PPh_3)_4$, Dioxane

(J. Med. Chem. 1996, 39, 217-223)

10) $R_{12}(H)C=CR_{13}R_{14}$, $Pd(OAc)_2$, Et_3N, DMF

(Org. React. 1982, 27, 345)

11) Tf_2O, THF

(J. Am. Chem. Soc. 1987, 109, 5478-5486)

12) $ArSnBu_3$, $Pd(PPh_3)_4$, Dioxane

(J. Am. Chem. Soc. 1987, 109, 5478-5486)

13) $KMnO_4$, Py, H_2O

(J. Med. Chem. 1996, 39, 217-223)

14) $NaOR_1$, THF

15) $NaSR_1$, THF

16) HNR_1R_{13}, THF

17) HONO, $NaBF_4$

(Adv. Fluorine Chem. 1965, 4, 1-30)

18) $Pd(OAc)_2$, NaH, DPPF, $PhCH_3$, R_1OH

(J. Org. Chem. 1997, 62, 5413-5418)

19) i. R_1X, Et_3N, CH_2Cl_2, ii. $R_{13}X$

20) $SOCl_2$, cat DMF

21) CH_2N_2, Et_2O

22) Ag_2O, Na_2CO_3, $Na_2S_2O_3$, H_2O

(Tetrahedron Lett. 1979, 2667)

23) AgO_2CPh, Et_3N, MeOH

(Org. Syn., 1970, 50, 77; J. Am. Chem. Soc. 1987, 109, 5432)

24) LiOH, THF-MeOH

25) $(EtO)_2P(O)CH_2CO_2R$, BuLi, THF

26) $MeO_2CCH(Br)=P(Ph)_3$, benzene

27) KOH or KOtBu

28) Base, $X(CH_2)_nCO_2R$

29) DPPA, Et_3N, toluene

(Synthesis 1985, 220)

30) HONO, H_2O

31) SO_2, CuCl, HCl, H_2O

(Synthesis 1969, 1-10, 6)

32) Lawesson's reagent, toluene

(Tetrahedron Asym. 1996, 7, 12, 3553)

33) R_2M, solvent

34) 30% H_2O_2, glacial CH_3CO_2H

(Helv. Chim. Acta. 1968, 349, 323)

35) triphosgene, CH_2Cl_2

(Tetrahedron Lett., 1996, 37, 8589)

36) i. $(EtO)_2P(O)CHLiSO_2Oi$-Pr, THF, ii. NaI

37) Ph_3PCH_3I, $NaCH_2S(O)CH_3$, DMSO

(Synthesis 1987, 498)

38) Br_2, $CHCl_3$ or other solvent

(Synthesis 1987, 498)

39) BuLi, Bu_3SnCl

40) $ClSO_2OTMS$, CCl_4

(Chem. Ber. 1995, 128, 575-580)

41) MeOH-HCl, reflux

42) LAH, Et_2O or $LiBH_4$, EtOH or BH_3-THF

(Tetrahedron Lett., 1996, 37, 8589)

43) MsCl, Et_3N, CH_2Cl_2

(Tetrahedron Lett., 1996, 37, 8589)

44) Na_2SO_3, H_2O

(Tetrahedron Lett., 1996, 37, 8589)

45) R_2R_4NH, Et_3N, CH_2Cl_2

46) R_2M, solvent

47) $CH_3NH(OCH_3)$, EDC, HOBt, DIEA, CH_2Cl_2 or DMF

(Tetrahedron Lett, 1981, 22, 3815)

48) MeLi, THF

49) mCPBA, CH_2Cl_2

50) HONO, Cu_2O, $Cu(NO_3)_2$, H_2O

(J. Org. Chem. 1977, 42, 2053)

51) R_1M, solvent

52) HONO, NaS(S)COEt, H_2O

(Org. Synth. 1947, 27, 81)

53) HSR_2 or HSR_4, CH_2Cl_2

54) i-BuOC(O)Cl, Et_3N, NH_3, THF

55) R_2R_4NH, CH_2Cl_2, $NaBH(OAc)_3$

56) R_2R_4NH, MeOH/CH_3CO_2H, $NaBH_3CN$

57) R_2OH, EDC, HOBt, DIEA, CH_2Cl_2 or DMF

58) R_2OH, HBTU, HOBt, DIEA, CH_2Cl_2 or DMF

59) R_2R_4NH, EDC, HOBt, DIEA, CH_2Cl_2 or DMF

60) R_2R_4NH, HBTU, HOBt, DIEA, CH_2Cl_2 or DMF

61) $POCl_3$, Py, CH_2Cl_2

62) R_2R_4NCO, solvent

63) $R_2OC(O)Cl$, Et_3N, solvent

64) R_2CO_2H, EDC or HBTU, HOBt, DIEA, CH_2Cl_2 or DMF

65) R_2X, Et_3N, solvent

66) $(CH_3S)_2C{=}N(CN)$, DMF, EtOH

(J. Med. Chem. 1994, 37, 57-66)

67) R_2SO_2Cl, Et_3N, CH_2Cl_2

68) R_2- or R_3- or R_4CHO, MeOH/CH_3CO_2H, $NaBH_3CN$

(Synthesis 1975, 135-146)

69) Boc(Tr)-D or L-CysOH, HBTU, HOBt, DIEA, CH_2Cl_2 or DMF

70) Boc(Tr)-D or L-CysH, $NaBH_3CN$, MeOH/CH_3CO_2H

(Synthesis 1975, 135-146)

71) S-Tr-N-Boc cysteinal, $ClCH_2CH_2Cl$ or THF, $NaBH(OAc)_3$

(J. Org. Chem. 1996, 61, 3849-3862)

72) TFA, CH_2Cl_2, Et_3SiH or (3:1:1) thioanisole/ethanedithiol/DMS

73) TFA, CH_2Cl_2

74) DPPA, Et_3N, toluene, $HOCH_2CH_2SiCH_3$

(Tetrahedron Lett. 1984, 25, 3515)

75) TBAF, THF

76) Base, TrSH or BnSH

77) Base, R_2X or R_4X

78) R_3NH_2, $MeOH/CH_3CO_2H$, $NaBH_3CN$

79) N_2H_4, KOH

80) $Pd_2(dba)_3$, P(o-tol)3, RNH_2, NaOtBu, Dioxane, R_1NH_2

(Tetrahedron Lett. 1996, 37, 7181-7184).

81) Cyanamide.

82) Fmoc-Cl, sodium bicarbonate.

83) BnCOCl, sodium carbonate.

84) AllylOCOCl, pyridine.

85) Benzyl bromide, base.

86) Oxalyl chloride, DMSO.

87) $RCONH_2$.

88) Carbonyldiimidazole, neutral solvents (e.g., DCM, DMF, THF, toluene).

89) Thiocarbonyldiimidazole, neutral solvents (e.g., DCM, DMF, THF, toluene).

90) Cyanogen bromide, neutral solvents (e.g., DCM, DMF, THF, toluene).

91) RCOCl, Triethylamine

92) $RNHNH_2$, EDC.

93) RO_2CCOCl, Et_3N, DCM.

94) MsOH, Pyridine (J. Het. Chem., 1980, 607.)

95) Base, neutral solvents (e.g., DCM, toluene, THF).

96) H_2NOR, EDC.

97) $RCSNH_2$.

98) RCOCHBrR, neutral solvents (e.g., DCM, DMF, THF, toluene), (Org. Proc. Prep. Intl., 1992, *24*, 127).

99) CH_2N_2, HCl. (Synthesis, 1993, 197).

100) NH2NHR, neutral solvents (e.g., DCM, DMF, THF, toluene).

101) RSO_2Cl, DMAP. (Tetrahedron Lett., 1993, *34*, 2749).

102) Et_3N, RX. (J. Org. Chem., 1990, *55*, 6037).

103) NOCl or Cl_2 (J. Org. Chem., 1990, *55*, 3916).

104) H_2NOH, neutral solvents (e.g., DCM, DMF, THF, toluene).

105) RCCR, neutral solvents (DCM, THF, Toluene).

106) RCHCHR, neutral solvents (DCM, THF, Toluene).

107) H_2NOH, HCl.

108) Thiocarbonyldiimidazole, SiO_2 or BF_3OEt_2. (J. Med. Chem., 1996, *39*, 5228).

109) Thiocarbonyldiimidazole, DBU or DBN. (J. Med. Chem., 1996, *39*, 5228).

110) HNO_2, HCl.

111) $ClCH_2CO_2Et$ (Org. Reactions, 1959, *10*,143).

112) Morpholine enamine (Eur. J. Med. Chem., 1982, *17*, 27).

113) RCOCHR'CN

114) $RCOCHR'CO_2Et$

115) Na_2SO_3

116) $H_2NCHRCO_2Et$

117) $EtO_2CCHRNCO$

118) $RCNHNH_2$.

119) $RCOCO_2H$, (J. Med. Chem., 1995, *38*, 3741).

120) RCHO, KOAc.

121) 2-Fluoronitrobenzene.

122) $SnCl_2$, EtOH, DMF.

123) RCHO, $NaBH_3CN$, HOAc.

124) NH_3, MeOH.

125) 2,4,6-$Me_3PhSO_2NH_2$.

126) Et_2NH, CH_2Cl_2

127) MeOC(O)Cl, Et_3N, CH_2Cl_2

128) R_2NH_2, EDC, HOBT, Et_3N, CH_2Cl_2

129) DBU, $PhCH_3$

130) $BocNHCH(CH_2STr)CH_2NH_2$, EDC, HOBT, Et_3N, CH_2Cl_2

131) $R_2NHCH_2CO_2Me$, HBTU, HOBT, Et_3N, CH_2Cl_2

132) $BocNHCH(CH_2STr)CH_2OMs$, LiHMDS, THF

133) $R_2NHCH_2CO_2Me$, $NaBH(OAc)_3$, $ClCH_2CH_2Cl$ or THF

134) $R_2NHCH_2CH(OEt)_2$, HBTU, HOBT, Et_3N, CH_2Cl_2

135) $NaBH(OAc)_3$, $ClCH_2CH_2Cl$ or THF, AcOH.

136) Piperidine, DMF.

137) $Pd(Ph_3P)_4$, Bu_3SnH.

138) RCO_2H, EDC, HOBT, Et_3N, DCM.

139) RNH_2, neutral solvents.

140) RCHO, $NaBH_3CN$, HOAc.

141) RNCO, solvent.

142) RCO_2H, EDC or HBTU, HOBt, DIEA, CH_2Cl_2 or DMF.

143) RCOCl, Triethylamine

144) RSO_2Cl, Et_3N, CH_2Cl_2.

145) $SnCl_2$, EtOH, DMF.

146) RNH_2, EDC, HOBt, DIEA, CH_2Cl_2 or DMF.

147) Dibromoethane, Et_3N, CH_2Cl_2

148) Oxalyl chloride, neutral solvents.

149) LiOH, THF-MeOH.

150) Carbonyldiimidazole, neutral solvents (e.g., DCM, DMF, THF, toluene).

151) RNH_2, Et_3N, CH_2Cl_2.

152) Base, RX.

153) DBU, $PhCH_3$

154) DPPA, Et_3N, toluene (Synthesis 1985, 220)

155) $SOCl_2$, cat DMF.

156) ArH, Lewis Acid ($AlCl_3$, $SnCl_4$, $TiCl_4$), CH_2Cl_2.

157) $H_2NCHRCO_2Et$, neutral solvents.

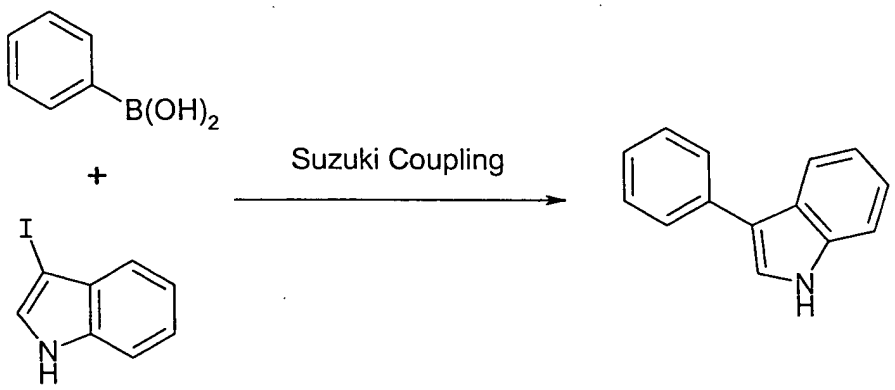
158) $BocHNCHRCO_2H$, EDC OR HBTU, HOBt, DIEA, CH_2Cl_2 or DMF.

159) TFA, CH_2Cl_2.

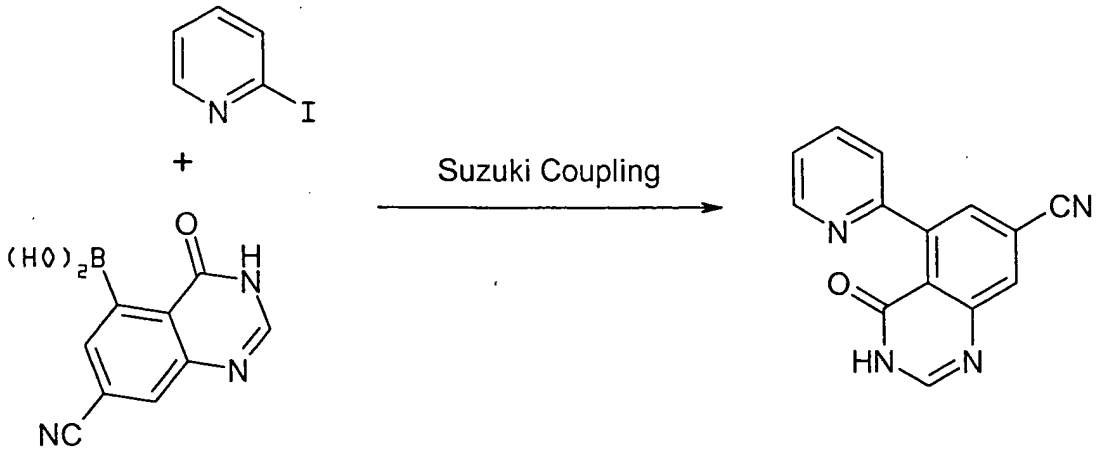
b. Illustrative preparation of aryl subunits

Ary subunits may be functionalized using a wide variety of reactions known to those in the art. The chemistry of aromatic and heteroaromatic rings is rich, and only a sampling of useful reactions can be presented here. A number of illustrative examples are shown below.

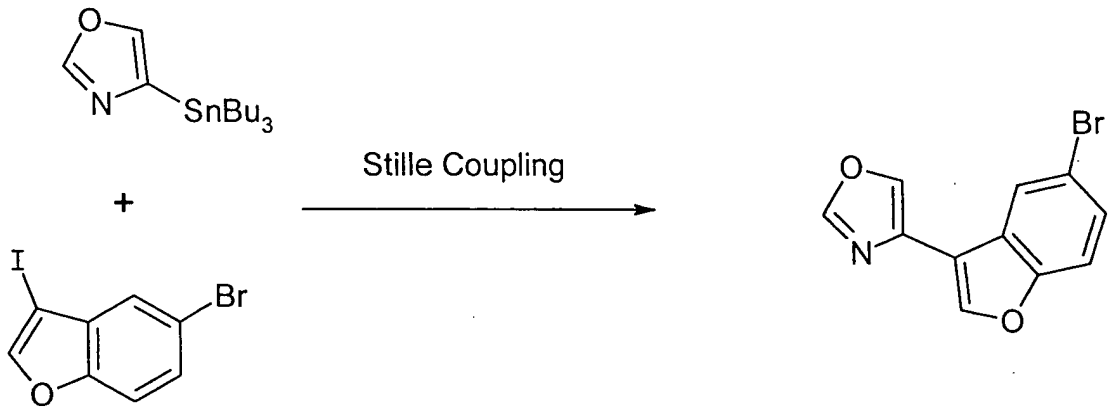
Suzuki Coupling No. 1:



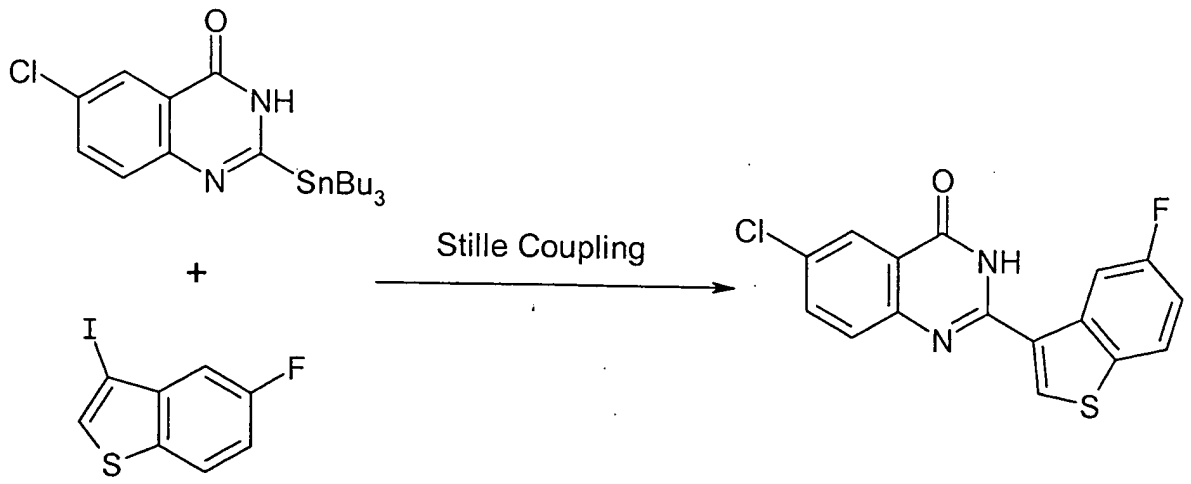
Suzuki Coupling No. 2:



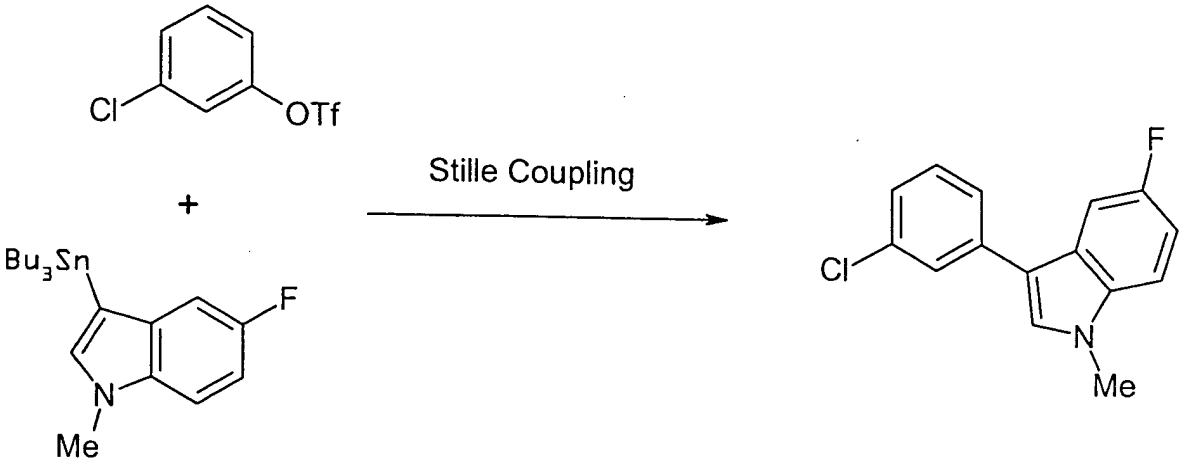
Stille Coupling No. 1:



Stille Coupling No. 2



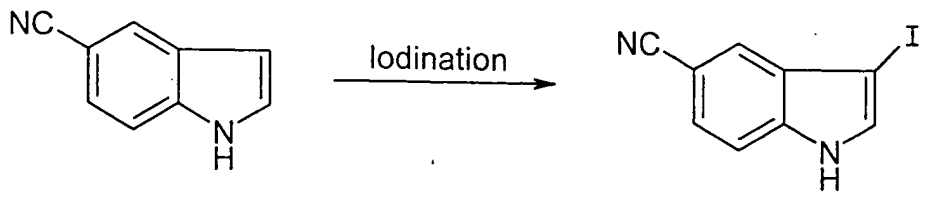
Stille Coupling No. 3



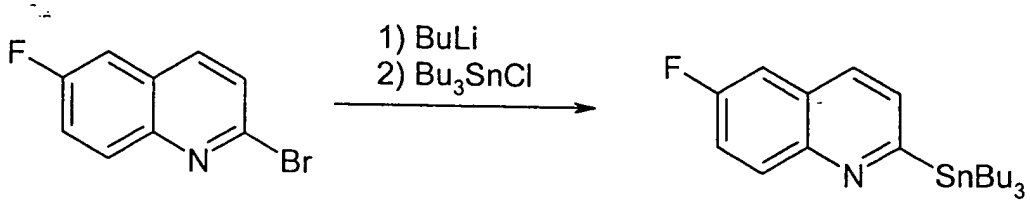
c. Illustrative preparation of coupling substrates

Members of the general classes of coupling substrates outlined above -- arylstannanes, arylboronic acids, aryl triflates and aryl halides -- are available from the parent heterocycles. In general, the transformations required to prepare a coupling substrate are reliable and amenable to scale-up. Illustrative examples are shown below.

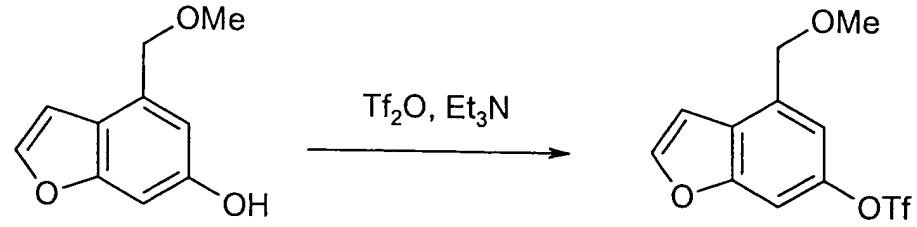
Preparation of an Aryl Iodide



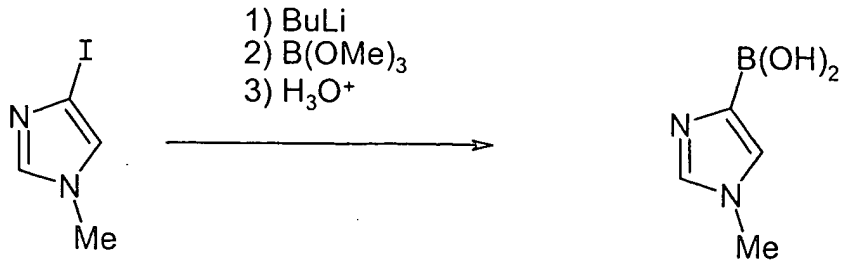
Preparation of an Aryl Stannane



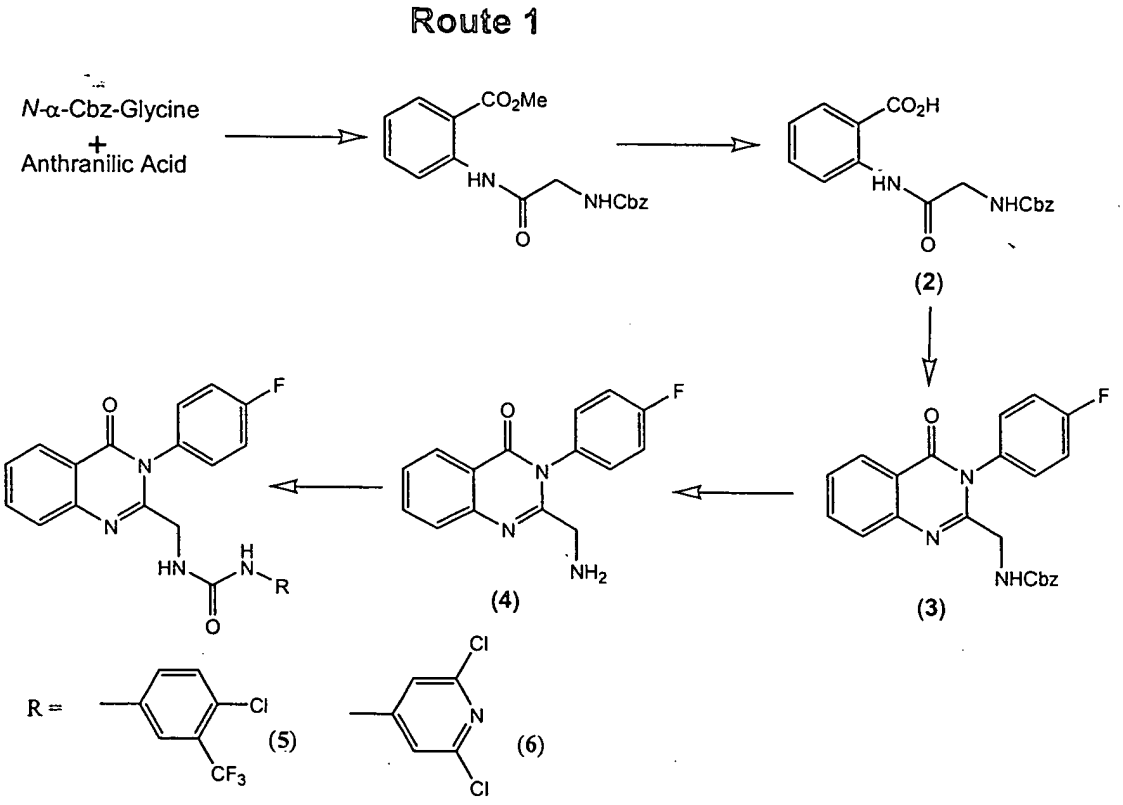
Preparation of an Aryl Triflate



Preparation of Aryl Boronic Acid



d. Exemplary Procedures for Preparing Subject Compounds



2-(2-Benzyloxycarbonylamino-acetylamino)-benzoic acid methyl ester (1)

To a solution of *N*-α-Cbz-glycine (2.0 g, 9.56 mmol) in tetrahydrofuran (20 mL) was added 1,1'-carbonyldiimidazole (1.64 g, 10.13 mmol), followed 2 h later by methyl anthranilate (1.28 g, 8.47 mmol). After stirring overnight the tetrahydrofuran was removed in *vacuo* and the residue was redissolved in ethyl acetate, washed with 10% aqueous hydrochloric acid (2x), saturated aqueous sodium hydrogen carbonate (2x), dried

($MgSO_4$) and concentrated to give an oily residue which was subjected to silica-gel column chromatography [eluent: Ethyl acetate:Hexane, 30:70, *v/v* -> Ethyl acetate] to give the *title* benzoate (**1**) (0.31 g, 11%) as a white solid:

δ (360 MHz; $CDCl_3$) 3.89 (s, 3H), 4.10 (d, 2H), 5.19 (s, 2H), 7.11 (t, 1H), 7.31-7.41 (m, 5H), 7.56 (t, 1H), 8.03 (d, 1H), 8.69 (d, 1H) and 11.54 (s, 1H).

2-(2-Benzyloxycarbonylamino-acetylamino)-benzoic acid (2)

Lithium hydroxide (55 mg, 1.32 mmol) in water (0.8 mL), was added to a solution of the benzoate (**1**) (300 mg, 0.88 mmol) in dioxane (4 mL). After stirring overnight, the reaction mixture was concentrated at 40 ^{o}C to give a viscous oil, which was treated with 10% aqueous hydrochloric acid (3 mL), with the resulting emulsion formed being extracted with ethyl acetate (4x). The organic phases were combined, dried ($MgSO_4$) and concentrated to give the *title* benzoic acid (**2**) quantitatively as a white solid:

δ (360 MHz; DMSO) 3.84 (d, 2H), 5.12 (s, 2H), 7.21 (t, 1H), 7.35-7.44 (m, 5H), 7.65 (t, 1H), 8.04 (t, 1H), 8.65 (d, 1H) and 11.74 (s, 1H).

[3-(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-ylmethyl]-carbamic acid benzyl ester (3)

To a solution of the benzoic acid (**2**) (180 mg, 0.55 mmol) in tetrahydrofuran (6 mL) was added 1,1'-carbonyldiimidazole (98 mg, 0.60 mmol). After a period of 2 h, 4-fluoroaniline (61 mg, 0.55 mmol) in tetrahydrofuran (1.5 mL) was added and the mixture was stirred at 75 °C overnight. The tetrahydrofuran was removed in *vacuo* and the residue was redissolved in ethyl acetate, washed with 10% aqueous hydrochloric acid, saturated aqueous sodium hydrogen carbonate, dried ($MgSO_4$) and concentrated to give an oily residue which was subjected to silica-gel column chromatography [eluent: Ethyl acetate:Hexane, 20:80, *v/v* -> 60:40, *v/v*] to give the *title* quinazolinone (**3**) (0.11 g, 50%) as a white solid:

δ (360 MHz; $CDCl_3$) 4.01 (d, 2H), 5.13 (s, 2H), 6.27 (s, 1H), 7.31-7.39 (m, 9H), 7.52 (t, 1H), 7.71 (d, 1H), 7.81 (t, 1H) and 8.28 (d, 1H).

2-Aminomethyl-3-(4-Fluoro-phenyl)-3*H*-quinazolin-4-one (4)

A stirring mixture of the quinazolinone (3) (63 mg, 0.16 mmol), 10% palladium on activated carbon (18 mg) and methanol (4.4 mL) was evacuated using an aspirator pump and filled with hydrogen. Once the starting material had been consumed as monitored by TLC analysis, the mixture was filtered through a Celite pad which was washed with methanol (2x) and concentrated to give the *title* amine (4) (40 mg, 93%) as a yellow solid:

δ (360 MHz; $CDCl_3$) 3.50 (s, 2H), 7.25 (apparent d, 4H), 7.50 (t, 1H), 7.75-7.82 (m, 2H) and 8.28 (d, 1H).

1-(4-Chloro-3-trifluoromethyl-phenyl)-3-[3-(4-fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-ylmethyl]-urea (5)

To a mixture of the amine (4) (14 mg, 0.05 mmol) in chloroform (0.5 mL), was added 3-trifluoromethylphenyl isocyanate (10 mg, 0.05 mmol). After stirring overnight the solvent was removed in *vacuo*, triturated with hexanes, filtered and dried to give the *title* urea (5) (24 mg, 98%) as a white solid:

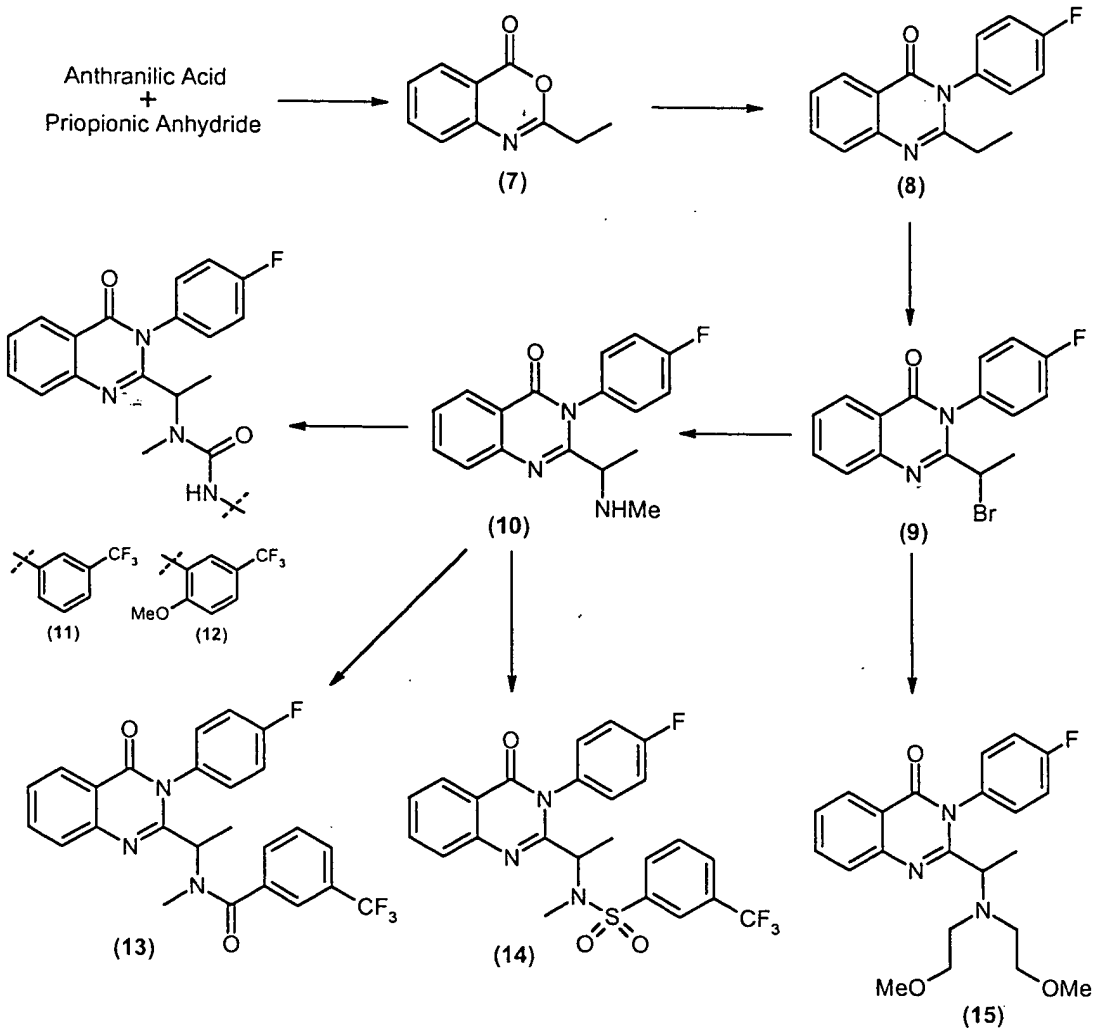
δ (360 MHz; $CDCl_3$) 4.08 (s, 2H), 6.52 (s, 1H), 7.16-7.38 (m, 4H), 7.48-7.53 (m, 2H), 7.59 (d, 1H), 7.66 (s 1H), 7.75 (t, 1H), 7.81 (s 1H) and 8.29 (d, 1H).

1-(2,6-Dichloro-pyridin-4-yl)-3-[3-(4-fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-ylmethyl]-urea (6)

To a mixture of the amine (4) (14 mg, 0.05 mmol) in chloroform (0.5 mL), was added 2,6-dichloropyridyl isocyanate (11 mg, 0.05 mmol). After stirring overnight the solvent was removed in *vacuo*, triturated with hexanes, filtered and dried to give the *title* urea (6) (22 mg, 96%) as a white solid:

δ (360 MHz MHz; $CDCl_3$) 4.09 (d, 2H), 6.41 (br s, 1H), 6.69 (dd, 1H), 6.75 (s, 1H), 6.87 (dd, 1H), 7.07 (t, 1H), 7.22-7.27 (m, 3H), 7.52 (t, 1H), 7.65 (d, 1H), 7.79 (t, 1H) and 8.29 (dd, 1H).

Route 2



2-Ethyl-3,1-[4H]benzoxazin-4-one (7)

Anthranilic acid (100 g, 0.73 mol) and propionic anhydride (420 mL) were heated to 140 ^{O}C for 3.5 h, in a flask equipped with Claisen-distillation head. The temperature was increased to 170-180 ^{O}C, and propionic acid and the propionic anhydride were removed under reduced pressure to give a brown solid, which was triturated with hexane, filtered and dried to give the *title* benzoxazine (7) (115.8 g, 91%) as an off white solid: δ (360 MHz; $CDCl_3$) 1.35 (t, 3H), 2.71 (q, 2H), 7.45-7.50 (m, 1H), 7.55 (d, 1H), 7.75-7.79 (m, 1H), 8.16 (dd, 1H).

2-Ethyl-3-(4'fluoro-phenyl)-quinazolin-4-one (8)

A solution of the benzoxazine (7) (5.0 g, 28.6 mmol) and 4-fluoroaniline (3.37 g, 30.3 mmol) in chloroform (12 mL) was heated to reflux overnight. After complete consumption of the benzoxazine (7) as monitored by TLC, the solvent was removed and the residual off white solid was taken up in ethylene glycol (8 mL) to which sodium hydroxide (51 mg) was added and the suspension was heated to 120-140^{O}C in a distillation apparatus. The water produced was distilled off during a 5 h heating period, and the dark solution allowed to cool to ambient temperature overnight. The pH was adjusted to 7-8 by addition of 3% aqueous hydrochloric acid, and the precipitate was filtered and dried to give the *title* quinazolinone **(8)** (4.42 g, 58%) as a white solid:

(360 MHz; $CDCl_3$) 1.23 (t, 3H), 2.44 (q, 2H), 7.24 (apparent d, 4H), 7.44-7.49 (m, 1H), 7.71-7.79 (m, 2H), 8.26 (dd, 1H).

2-(1'Bromoethyl)-3-(4''fluorophenyl)-quinazolin-4-one (9)

To a stirring mixture of the quinazolinone **(8)** (94.0 g, 0.35 mol), anhydrous sodium acetate (35.2 g, 0.43 mol) and glacial acetic acid (448 mL) at 40^{O}C under a nitrogen atmosphere, was added dropwise a solution of bromine (76.8 g, 0.48 mol) in glacial acetic acid (243 mL) whilst maintaining the mixture at approximately 40^{O}C over a 5 h period. The mixture was poured onto water (5.26 L), stirred for 1 h and filtered. The filter cake was washed with warm water (2.8 L) and dried to give the *title* bromide **(9)** (117.2 g, 96%) as a white solid:

δ (360 MHz; $CDCl_3$) 1.99 (d, 3H), 4.48 (q, 1H), 7.07-7.25 (m, 3H), 7.44-7.52 (m, 2H), 7.73-7.75 (m, 2H), 8.21 (d, 1H).

2-[(1'-Methylamino)-ethyl]-3-(4''-fluorophenyl)-quinazolin-4-one (10)

A 8.03M solution of methylamine in ethanol (923 mL, 7.41 mol), was added to the bromide **(9)** (117.0 g, 0.34 mol) and the suspension was warmed to 40^{O}C. After 1 h the starting material was fully consumed as monitored by TLC and the solvent and excess methylamine were removed under reduced pressure. The residue was taken up in

dichloromethane (1.2 L), stirred for 1 h at ambient temperature, with the precipitated methylammonium hydrochloride being filtered off, and washed with dichloromethane (117 mL). The combined organic phases were concentrated and dried ($MgSO_4$) to give the *title* amine (**10**) quantitatively as a white solid:

δ (360 MHz; $CDCl_3$) 1.30 (d, 3H), 2.33 (s, 3H), 3.38 (q, 1H), 7.26-7.34 (m, 4H), 7.51-7.55 (m, 1H), 7.76-7.85 (m, 2H), 8.31 (dd, 1H).

2-{1'-[N-(3''-trifluoromethylphenylcarbamoyl)-N-methyl-amino]ethyl}-3-(4'''fluorophenyl)-quinazolin-4-one (11)

To a solution of the amine (**10**) (18.9 g, 0.065 mol) and chloroform (150 mL) was added a solution of 3-trifluoromethylphenyl isocyanate (11.9 g, 0.065 mol) in chloroform (39 mL) whilst maintaining the temperature below 30°C. The mixture was stirred at ambient temperature for 1 h, and the precipitate formed was filtered, washed with chloroform and dried to give the *title* urea (**11**) (20.2 g, 65%) as an off white solid:

δ (360 MHz; $CDCl_3$) 1.55 (d, 3H), 2.92 (s, 3H), 5.16 (q, 1H), 7.09 (br s, 1H), 7.18-7.23 (m, 1H), 7.29-7.30 (m, 3H), 7.34-7.43 (m, 2H), 7.55-7.58 (m, 3H), 7.80-7.88 (m, 2H), 8.31 (d, 1H).

2-{1'-[N-(2''-methoxy-5''-trifluoromethylphenylcarbamoyl)-N-methyl-amino]ethyl}-3-(4'''fluorophenyl)-quinazolin-4-one (12)

A solution of 2-methoxy-5-trifluoromethylaniline (42 mg, 0.22 mmol) in ethyl acetate (1.2 mL) was added to a solution of triphosgene (80 mg, 0.27 mmol) in ethyl acetate (1.2 mL), catalytic amounts of charcoal were added and the mixture heated to reflux for 2 h. After cooling to ambient temperature the solvent was evaporated under reduced pressure, with the resulting residue being taken up in chloroform (1.2 mL). A solution of the amine (**10**) (65 mg, 0.22 mmol) in chloroform (1.2 mL) was added and the mixture stirred at ambient temperature until the starting material was fully consumed. The solvent was evaporated and the crude product was subjected to silica-gel column chromatography [eluent: Dichloromethane -> Ethyl acetate], with the *title* urea (**12**) (98 mg, 87%) being isolated as white solid:

δ (360 MHz; $CDCl_3$) 1.40 (d, 3H), 2.88 (s, 3H), 3.83 (s, 3H), 5.25 (q, 1H), 6.80 (d, 1H), 6.97-7.03 (m, 2H), 7.13-7.18 (m, 3H), 7.20-7.25 (m, 1H), 7.40-7.45 (m, 1H), 7.65-7.74 (m, 2H), 8.18-8.22 (m, 2H).

N-{1-[3-(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-yl]-ethyl}-N-methyl-3-trifluoromethyl-benzamide (13)

To a solution of the amine (**10**) (50 mg, 0.17 mmol) and *N,N*-diisopropylethylamine (0.03 mL, 0.17 mmol) in dichloromethane (0.5 mL) was added 3-trifluoromethylbenzoyl chloride (0.03 mL, 0.19 mmol). The mixture was stirred overnight at ambient temperature under a nitrogen atmosphere. After 12 h a few crystals of 4-dimethylaminopyridine were added, and stirring was continued for another 12 h. The mixture was quenched with water, washed with saturated aqueous sodium hydrogen carbonate, dried ($MgSO_4$), filtered and concentrated to furnish the *title* benzamide (**13**) quantitatively as a white foam:

δ (360 MHz; $CDCl_3$) 1.53 (d, 3H), 3.16 (s, 3H), 5.36 (q, 1H), 7.24-7.36 (m, 3H), 7.45-7.56 (m, 3H), 7.64-7.82 (m, 5H), 8.29 (br d, 1H).

N-{1-[3-(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-yl]-ethyl}-N-methyl-3-trifluoromethyl-benzenesulfonamide (14)

To a solution of the amine (**10**) (50 mg, 0.17 mmol) and *N,N*-diisopropylethylamine (0.03 mL, 0.17 mmol) in dichloromethane (0.5 mL) was added 3-trifluoromethylphenylsulfonyl chloride (0.03 mL, 0.19 mmol), and the mixture was stirred at ambient temperature under a nitrogen atmosphere. After 12 h a few crystals of 4-dimethylaminopyridine were added, and stirring was continued for another 12 h. The mixture was quenched with water, washed with saturated aqueous sodium hydrogen carbonate, dried ($MgSO_4$), filtered and concentrated to furnished the *title* sulfonamide (**14**) quantitatively as a white foam:

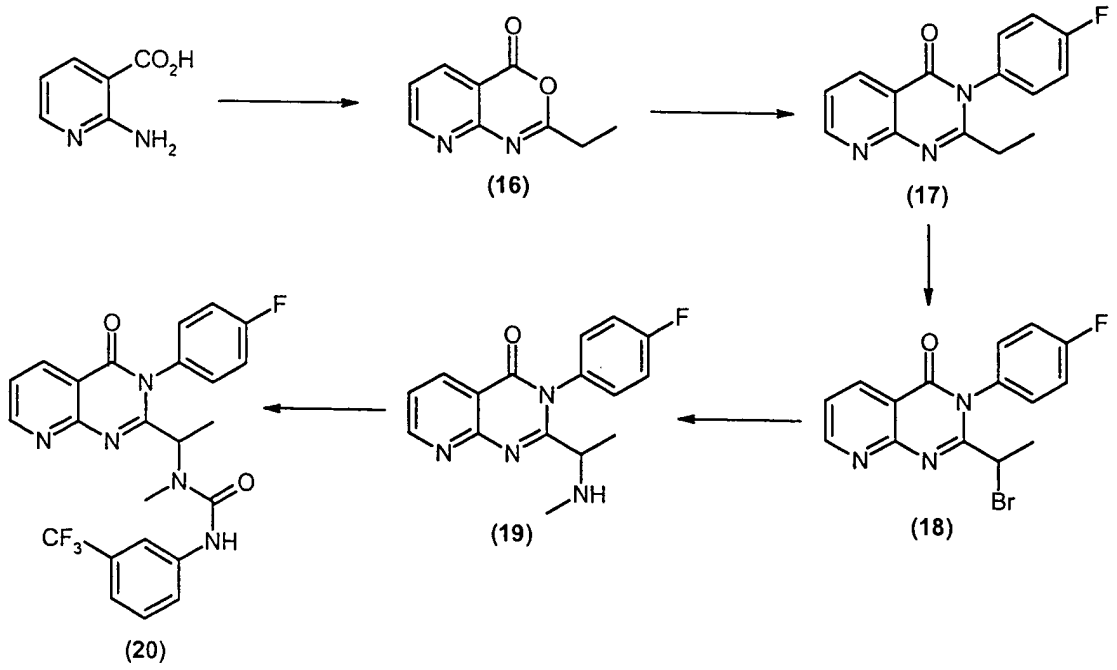
δ (360 MHz; $CDCl_3$) 1.28 (d, 3H), 3.16 (s, 3H), 4.92 (q, 1H), 7.18-7.30 (m, 2H), 7.35-7.50 (m, 5H), 7.62 (br d, 1H), 7.70 (dt, 1H), 7.79 (br d, 1H), 7.87 (br s, 1H), 8.25 (dd, 1H).

2-{1-[Bis-(2-methoxy-ethyl)-amino]-ethyl}-3-(4-fluoro-phenyl)-3*H*-quinazolin-4-one (15)

A solution of the bromide (**9**) (71.5 mg, 0.21 mmol), tetrahydrofuran (1.0 mL) and *bis*-(2-methoxyethyl)amine (0.33 mL, 2.26 mmol) was heated to 70°C. After 6.5 h, the reaction was partly concentrated, and ethyl acetate (1 mL) and water (1 mL) were added with the mixture being stirred vigorously for 1 h. The organic layer was separated, washed with water (2 x 1 mL), dried ($MgSO_4$), filtered and concentrated to give the *title* amine (**15**) (61 mg, 74%) as an oil:

δ (360 MHz; $CDCl_3$) 1.35 (d, 3H), 2.51-2.56 (m, 2H), 2.74-2.83 (m, 2H), 2.98-3.11 (m, 4H), 3.12 (s, 6H), 3.79 (q, 1H), 7.13-7.27 (m, 3H), 7.45-7.54 (m, 2H), 7.74-7.76 (m, 2H), 8.26 (d, 1H).

Route 3



2-Ethyl-pyrido[2,3-*d*][1,3]oxazin-4-one (16)

A suspension of 2-aminonicotinic acid (1.0 g 7.24 mmol) in propionic anhydride (10 mL) of was heated to 120°C for 1 h under nitrogen atmosphere. The temperature was increased to 167°C to remove the propionic acid and propionic anhydride under reduced pressure. The resulting brown solid was triturated with hexane and dried to give the *title* pyrido-oxazine (**16**) (1.10 g, 86%) as a white solid:

δ (360 MHz; $CDCl_3$) 1.43 (t, 3H), 2.82 (q, 2H), 7.48 (dd, 1H), 8.53 (dd, 1H), 8.98 (d, 1H).

2-Ethyl-3-(4-fluoro-phenyl)-3*H*-pyrido[2,3-*d*]pyrimidin-4-one (17)

A suspension of the pyrido-oxazine (**16**) (0.5 g, 2.84 mmol) and 4-fluoroaniline (0.32 g, 2.84 mmol) in toluene (21 mL) was heated to reflux overnight. The solution was concentrated to dryness under reduced pressure and ethylene glycol (2 mL) and sodium hydroxide (5 mg) were added. The mixture was heated again to 120°C for 4 h, and the ethylene glycol was distilled off under reduced pressure. The resulting dark brown solid was subjected to silica-gel column chromatography [eluent: Ethyl acetate:Hexane, 20:80, *v/v* -> 100:0, *v/v*] to give the *title* pyrido-pyrimidinone (**17**) (244 mg, 32%) as a beige solid:

δ (360 MHz; $CDCl_3$) 1.31 (t, 3H), 2.49 (q, 2H), 7.25-7.28 (m, 4H), 7.44 (dd, 1H), 8.59 (dd, 1H), 9.00 (dd, 1H).

2-(1-Bromo-ethyl)-3-(4-fluoro-phenyl)-3*H*-pyrido[2,3-*d*]pyrimidin-4-one (18)

A suspension of the quinazolinone (**17**) (50.0 g, 0.56 mmol) and anhydrous sodium acetate (56 mg) in glacial acetic acid (1.0 mL) was heated to 40°C and treated dropwise with a solution of bromine (90 mg, 0.56 mmol) in glacial acetic acid (5 mL). Once the starting material was fully consumed by TLC analysis, the mixture was poured onto water (5 mL), basified with saturated aqueous sodium hydrogen carbonate and extracted with *tert*-butylmethyl ether. The combined organic phases were dried ($MgSO_4$), filtered, and concentrated to a solid which was purified by silica-gel column chromatography [eluent: Ethyl acetate:Hexane, 50:50, *v/v*] to give the *title* bromide (**18**) (166 mg, 85%) as an off white solid:

δ (360 MHz; $CDCl_3$) 1.21 (d, 3H), 4.58 (q, 1H), 7.15-7.34 (m, 3H), 7.51 (dd, 1H), 7.55-7.60 (m, 1H), 8.62 (dd, 1H), 9.05 (dd, 1H).

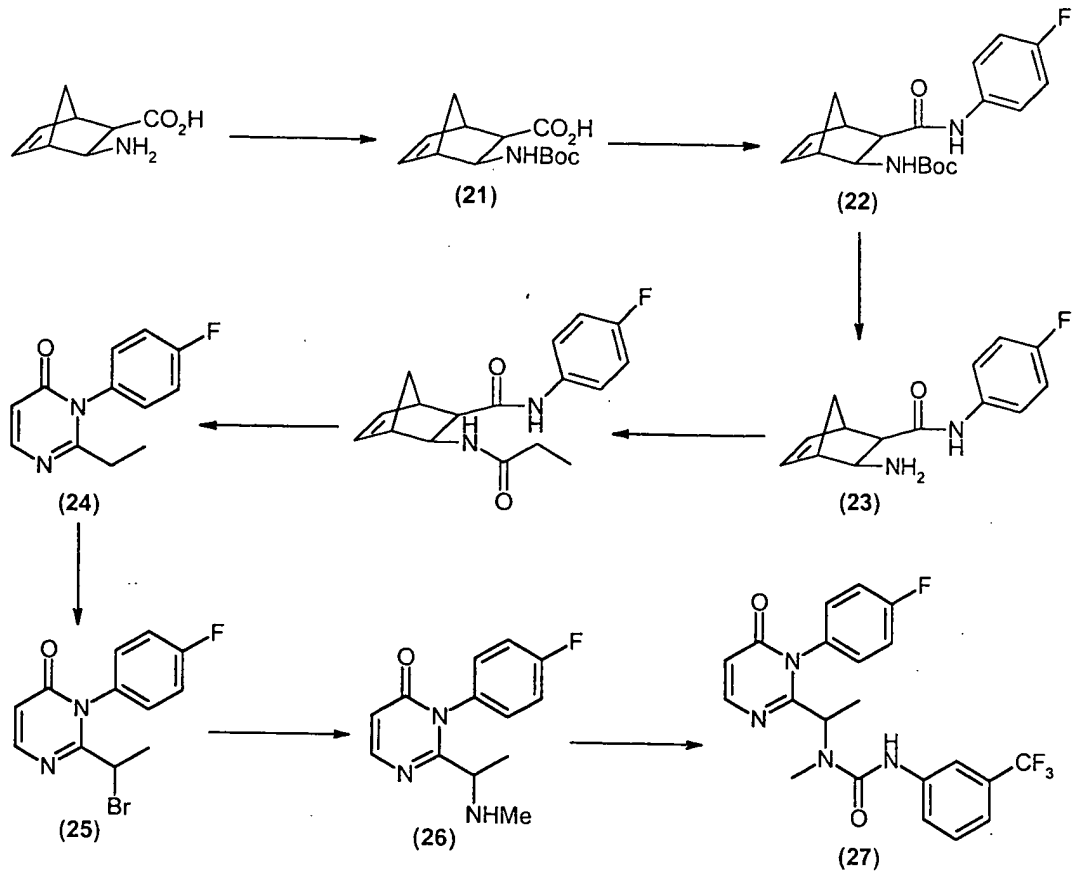
3-(4-Fluoro-phenyl)-2-(1-methylamino-ethyl)-3*H*-pyrido[2,3-*d*]pyrimidin-4-one (19)

A suspension of the bromide (**18**) (80 mg, 0.23 mmol) and 8.03M solution of methylamine in ethanol (0.64 mL) was heated to 40°C for 3 h, and stirred at ambient temperature overnight. The resulting suspension was concentrated to dryness under reduced pressure, and the resulting light brown solid was taken up in dichloromethane and filtered through a silica-gel plug to give the *title* amine (**19**) (57 mg, 83%) as a beige solid (The crude material was directly taken through the next step without further purification).

1-{1-[3[(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-pyrido[2,3-*d*]pyrimidin-2-yl]-ethyl}-1-methyl-3-(3-trifluoromethyl-phenyl)-urea (20)

To a solution of the amine (**19**) (30 mg, 0.1 mmol) in diclhoromethane (0.2 mL) was added 3-trifluoromethylphenyl isocyanate (19 mg, 0.1 mmol). The mixture was stirred at ambient temperature overnight, after which, the precipitate was filtered, washed with dichloromethane and dried to give the title urea (**20**) (6 mg, 13%) as a white solid: δ (360 MHz; $CDCl_3$) 2.17 (d, 3H), 3.15 (s, 3H), 5.23 (q, 1H), 7.15 (br s, 1H), 7.22-7.31 (m, 4H), 7.38 (t, 1H), 7.49 (d, 1H), 7.50-7.56 (m, 2H), 7.62 (br s, 1H), 8.62 (dd, 1H), 9.00 (d, 1H).

Route 4



3-Exo-tert-butoxycarbonylamino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid (21)

A solution of 3-exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid (1.0 g, 6.5 mmol), 1N aqueous sodium hydroxide (6.5 mL) and dioxane (6.5 mL) was cooled on an ice-bath, and di-*tert*-butyl dicarbonate (3.0 g, 13.8 mmol) was added with stirring continuing for 10 minutes at $0^{o}C$ and 6 h at ambient temperature. The solvent was partially evaporated and the pH was adjusted to 1-2 by addition of 1N aqueous potassium hydrogen sulfate (~8 mL). The aqueous phase was extracted with ethyl acetate (3 x 10 mL), and the combined organic phases were dried ($MgSO_4$), and concentrated to give the *title* acid (**21**) (1.41 g, 85%) as a white solid:

δ (360 MHz; $CDCl_3$) 1.70 (s, 9H), 1.89 (m, 1H), 2.31 (m, 1H), 2.84 (m, 1H), 2.98 (m, 1H), 3.22 (m, 1H), 4.20 (m, 1H), 6.44 (m, 2H), 7.21 (m, 1H).

[3-Exo-(4-fluoro-phenylcarbamoyl)-bicyclo[2.2.1]hept-5-en-2-exo-yl]-carbamic acid tert-butyl ester (22)

A solution of the acid (**21**) (1.4 g, 5.5 mmol) in anhydrous tetrahydrofuran (20 mL) was cooled to –10 ^{O}C and treated dropwise with triethylamine (0.77 mL, 5.5 mmol), followed by isobutyl chloroformate (0.72 mL, 5.5 mmol). The resulting suspension was stirred for 10 minutes at –10 ^{O}C before a solution of 4-fluoroaniline (0.53 mL, 5.5 mmol) in anhydrous tetrahydrofuran (5 mL) was added, with the temperature being kept between –8 and –17 ^{O}C for 5 h and mixture allowed to warm to ambient temperature overnight. The triethylaminehydrochloride was filtered off, washed with tetrahydrofuran, and the solvent was evaporated to dryness, with the resulting solids being washed with water (20 mL) and dried to give the title carbamic acid ester (**22**) quantitatively as a white solid:
δ (360 MHz; $CDCl_3$) 1.24 (s, 9H), 1.66 (d, 1H), 2.06 (d, 1H), 2.45 (d, 1H), 2.71 (br s, 1H), 3.08 (br s, 1H), 3.93-4.06 (m, 1H), 4.96 (d, 1H), 6.18-6.24 (m, 2H), 6.94-7.02 (m, 2H), 7.48-7.50 (m, 2H), 7.98 (br s, 1H).

3-Exo-amino-bicyclo[2.2.1]hept-5-ene-2-exo-carboxylic acid (4-fluoro-phenyl)-amide (23)

To a solution of the carbamic acid ester (**22**) (2.24 g) in dichloromethane (50 mL) was added trifluoroacetic acid (5 mL). The solution was stirred at ambient temperature for 4 h before water (50 mL) was added. The aqueous phase was separated and washed with dichloromethane (25mL), filtered and basified with 2N sodium hydroxide. The basified aqueous phase was extracted with dichloromethane (2 x 25 mL), and the combined organic phases were washed with brine (10 mL), dried ($MgSO_4$), and concentrated to give the title amine (**23**) (870 mg, (64% over 2 steps from (**21**)) as a solid:
δ (360 MHz; $CDCl_3$) 1.59 (d, 1H), 1.71 (br d, 2H), 2.13 (d, 1H), 2.37 (d, 1H), 2.59 (br s, 1H), 3.09 (br s, 1H), 3.24 (d, 1H), 6.20 (br s, 2H), 6.99 (t, 2H), 7.47-7.51 (m, 2H), 8.54 (br s, 1H).

2-Ethyl-3-(4-fluoro-phenyl)-3H-pyrimidin-4-one (24)

A solution of the amine (**23**) (43 mg, 0.17 mmol) in triethylorthopropionate (0.5 mL, 2.38 mmol) was heated to 100°C for 26 h, after which the excess solvent and volatile co-products were evaporated under reduced pressure to give the *title* pyrimidinone (**24**) (41 mg, quant.) as a crystalline solid:

δ (360 MHz; $CDCl_3$) 1.17 (t, 3H), 2.37 (q, 2H), 6.45 (d, 1H), 7.17-7.27 (m, 4H), 7.93 (d, 1H).

2-(1-Bromo-ethyl)-3-(4-fluoro-phenyl)-3H-pyrimidin-4-one (25)

A solution of the pyrimidinone (**24**) (150 mg, 0.69 mmol) and sodium acetate (140 mg) in glacial acetic acid (1.5 mL) was warmed to 40°C and treated dropwise with a preformed bromine solution (0.7 mL bromine in 10 mL of glacial acetic acid) (0.55 mL, 0.69 mmol). After 2 h, water (20 mL) was added to the mixture which was subsequently basified with potassium carbonate, and extracted with dichloromethane (2 x 10 mL). The combined organic phases were washed with water (10 mL), dried ($MgSO_4$), and concentrated to give the *title* bromide (**25**) (196 mg, 96%) as a pale brown oil:

δ (360 MHz; $CDCl_3$) 1.95 (d, 3H), 4.46 (q, 1H), 6.50 (d, 1H), 7.12-7.16 (m, 1H), 7.20-7.28 (m, 2H), 7.43-7.50 (m, 1H), 8.81 (d, 1H).

3-(4-Fluoro-phenyl)-2-(1-methylamino-ethyl)-3H-pyrimidin-4-one (26)

The bromide (**25**) (190 mg, 0.64 mmol) was dissolved in a 33% solution of methylamine in ethanol (5 mL) and stirred at ambient temperature. After 3.5 h, the solvent and excess methylamine were evaporated under reduced pressure and residue was partitioned between water (10 mL) and dichloromethane (10 mL). The organic layer was washed again with water (10 mL), dried ($MgSO_4$), and concentrated to give the title amine (**26**) (120 mg, 76%) as a yellow oil:

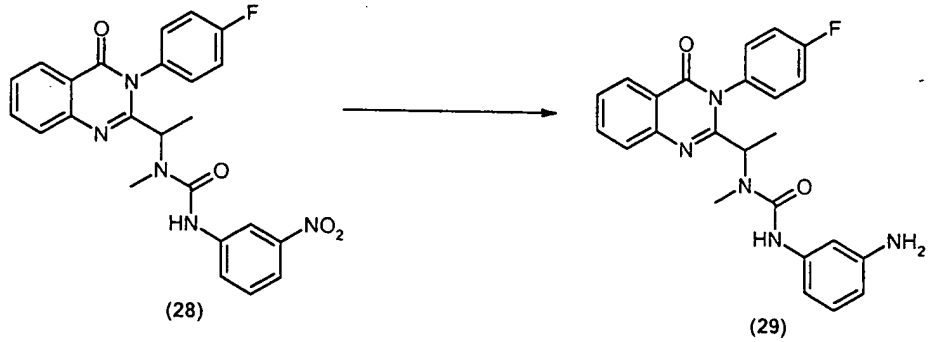
δ (360 MHz; $CDCl_3$) 1.20 (d, 3H), 2.23 (s, 3H), 3.26 (q, 1H), 6.44 (d, 1H), 7.16-7.33 (m, 4H), 7.98 (d, 1H).

1-{1-[1-(4-Fluoro-phenyl)-6-oxo-1,6-dihydro-pyrimidin-2-yl]-ethyl}-1-methyl-3-(3-trifluoromethyl-phenyl)-urea (27)

The amine (**26**) (12 mg, 0.05 mmol) and 3-trifluoromethylphenyl isocyanate (7.5 L, 0.05 mmol) were dissolved in chloroform (0.2 mL) and stirred at ambient temperature. After complete conversion of the starting material by TLC analysis, the solvent was removed and the resulting residue was subjected to silica-gel column chromatography [eluent: Dichloromethane] to give the *title* urea (**27**) (14 mg, quant.) as a solid:

δ (360 MHz; $CDCl_3$) 1.61 (m, 3H), 2.87 (s, 3H), 5.04 (m, 1H), 6.41 (m, 1H), 6.64 (s, 1H), 7.11-7.48 (m, 8H), 7.86 (m, 1H).

Nitro Reduction to Anilines

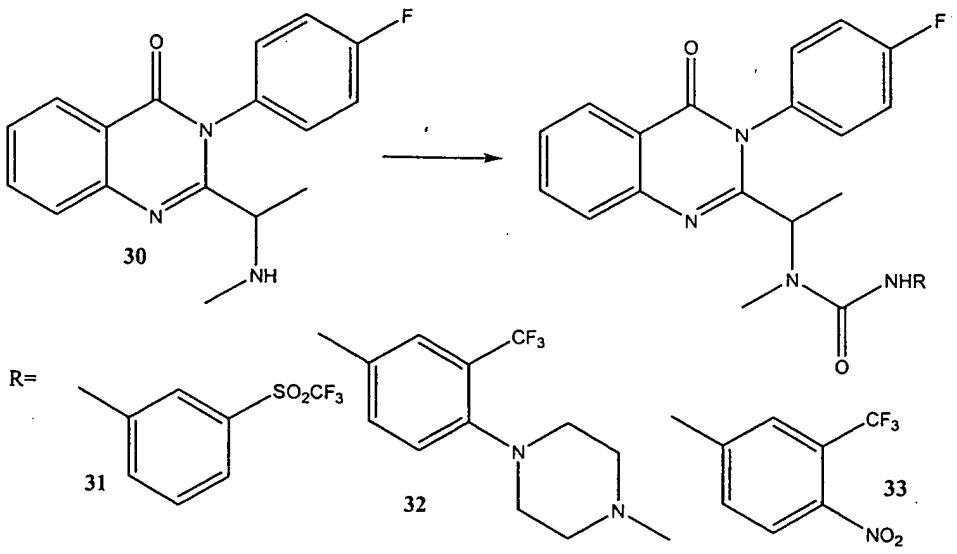


2-{1'-[N-(3''-amino-carbamoyl)-N-methyl-amino]ethyl}-3-(4'''fluorophenyl)-quinazolin-4-one (29)

A solution of 2-{1'-[N-(3''-nitro-carbamoyl)-N-methyl-amino]ethyl}-3-(4'''fluorophenyl)-quinazolin-4-one (**28**) (168 mg, 0.036 mmol), dichloromethane (5 mL) and ethyl acetate (5 mL) was stirred under a hydrogen atmosphere in the presence of 10% activated palladium on charcoal until the starting material was fully consumed. The mixture was filtered through celite and concentrated to dryness to give the *title* aniline (**29**) (100 mg, 65%):

δ (360 MHz; $CDCl_3$) 1.48 (d, 3H), 2.88 (s, 3H), 5.17 (q, 1H), 6.39 (br s, 1H), 6.50 (br s, 1H), 6.92 (br s, 1H), 6.60 (br s, 1H), 7.02 (t, 1H), 7.15-7.33 (m, 4H), 7.49-7.55 (m, 1H), 7.75-7.82 (m, 2H), 8.27 (d, 1H).

An *in situ* method for derivatizing the pendant amine was developed, as exemplified by the following protocol:



1-{1-[3-(Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-yl]-ethyl}-1-methyl-3-(3-trifluoromethanesulfonyl-phenyl)-urea (31)

A solution of triphosgene (160 mg, 0.54 mmol) and ethyl acetate (2.5 mL) was added to a solution of (3-aminophenyl)trifluoromethyl sulfone (100 mg, 0.44 mmol) and ethyl acetate (2.5 mL). After stirring for 5 min the mixture was refluxed until it went clear. The mixture was concentrated and redissolved in chloroform (2.5 mL) to which a solution of the amine (**30**) (130 mg, 0.44 mmol) in chloroform (2.5 mL) was added. Once all of the starting material had been consumed by TLC, the solvent was removed in *vacuo* and subjected to silica gel column chromatography using dichloromethane then ethyl acetate as eluent to give the *title urea* (**31**) (226 mg, 94%) as a solid.

δ_H (360 MHz; $CDCl_3$) 1.43 (d, 3H), 2.88 (s, 3H), 4.98 (br s, 1H), 7.08-7.14 (m, 1H), 7.16-7.31 (m, 3H), 7.45-7.54 (m, 3H), 7.60 (d, 1H), 7.69-7.80 (m, 3H), 7.96 (d, 1H) and 8.22 (d, 1H).

1-{1-[3-(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-yl]-ethyl}-3-(3-trifluoromethyl -4-nitro-phenyl)-1-methyl-urea (32)

A solution of triphosgene (80 mg, 0.27 mmol) and ethyl acetate (1.2 mL) was added to a solution of 4-nitro-3-trifluoromethylaniline (45 mg, 0.22 mmol) and ethyl acetate (1.2 mL). After stirring for 5 min the mixture was refluxed until it went clear. The mixture was concentrated and redissolved in chloroform (1.2 mL) to which a solution of the amine (**30**) (65 mg, 0.22 mmol) in chloroform (1.2 mL) was added. Once all of the starting material had been consumed by TLC, the solvent was removed in *vacuo* and subjected to silica gel column chromatography using dichloromethane then ethyl acetate as eluent to give the *title urea* (**32**) (112 mg, 97%) as a solid.

δ_H (360 MHz; $CDCl_3$) 1.45 (d, 3H), 2.80 (s, 3H), 4.90 (br s, 1H), 7.05-7.25 (m, 4H), 7.48 (t, 1H), 7.63-7.79 (m, 4H), 7.86 (d, 1H) and 8.19 (d, 1H).

1-{1-[3-(4-Fluoro-phenyl)-4-oxo-3,4-dihydro-quinazolin-2-yl]-ethyl}-3-[3-trifluoromethyl-4-(4-methyl)-piperazin-1-yl)-phenyl]-1-methyl-urea (33)

A solution of triphosgene (80 mg, 0.27 mmol) and ethyl acetate (1.2 mL) was added to a solution of 4-(*N*-methylpiperazine)-3-trifluoromethyl aniline (56 mg, 0.22 mmol) and ethyl acetate (1.2 mL). After stirring for 5 min the mixture was refluxed until it went clear. The mixture was concentrated and redissolved in chloroform (1.2 mL) to which a solution of the amine (**30**) (65 mg, 0.22 mmol) in chloroform (1.2 mL) was added. Once all of the starting material had been consumed by TLC, the solvent was removed in *vacuo* and subjected to silica gel column chromatography using ethyl acetate then 5-10% methanol in dichloromethane as eluent to give the *title urea* (**33**) (25 mg, 20%) as a solid.

δ_H (360 MHz; $CDCl_3$) 1.38 (d, 3H), 1.97 (br s, 2H), 2.26 (s, 3H), 2.47 (br s, 2H), 2.72-2.86 (m, 7H), 4.92-5.05 (m, 1H), 6.80-6.93 (br s, 1H), 7.08 (t, 1H), 7.12-7.27 (m, 4H), 7.34 (s, 1H), 7.41 (t, 2H), 7.63-7.74 (m, 2H) and 8.16 (d, 1H).

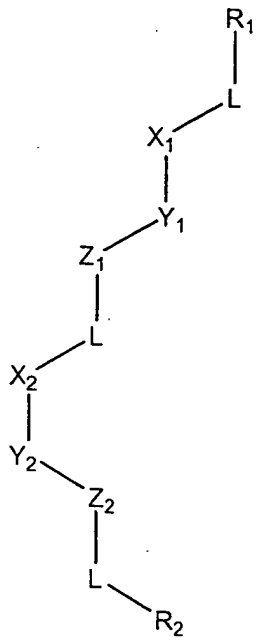
All of the references cited above are hereby incorporated by reference herein.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

We claim:

1. A method for inhibiting an altered growth state of a cell having a *hedgehog* gain-of-function phenotype, comprising contacting the cell with a *hedgehog* antagonist in a sufficient amount to inhibit the altered growth state, wherein the *hedgehog* antagonist is a organic molecule represented in the general formula (I):



Formula I

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X_1 and X_2 are selected, independently, from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-N(R_8)-, -ON(R_8)-, a heterocycle, or a direct bond between L and Y_1 or Y_2, respectively;

Y_1 and Y_2 are selected, independently, from -C(=O)-, -C(=S)-, -S(O_2)-, -S(O)-, -C(=NCN)-, -P(=O)(OR_2)-, a heteroaromatic group, or a direct bond between X_1 and Z_1 or X_2 and Z_2, respectively;

Z_1 and Z_2 are selected, independently, from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y_1 or Y_2, respectively, and L;

R_8, independently for each occurrence, represents H, lower alkyl, -(CH_2)$_n$aryl (substituted or unsubstituted), -(CH_2)$_n$heteroaryl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

2. The method of claim 1, wherein R_1 represents a substituted or unsubstituted aryl or heteroaryl group.

3. The method of claim 1, wherein at least one of X_1-Y_1-Z_1 and X_2-Y_2-Z_2 taken together represents a urea or an amide.

4. The method of claim 1, wherein R_1 represents either a fused cycloalkyl-aryl or cycloalkyl-heteroaryl system.

5. The method of claim 1, wherein the *hedgehog* antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 mM or less.

6. The method of claim 1, wherein the hedgehog antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 μM or less.

7. The method of claim 1, wherein the hedgehog antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 nM or less.

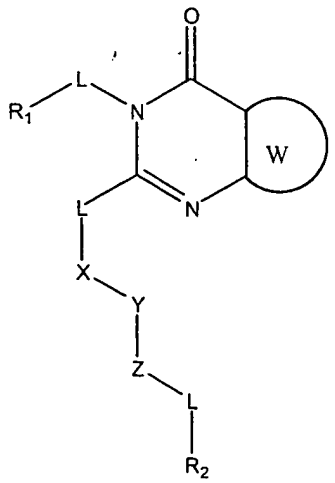
8. The method of claim 1, wherein the cell is contacted with the *hedgehog* antagonist *in vitro*.

9. The method of claim 1, wherein the cell is contacted with the *hedgehog* antagonist *in vivo*.

10. The method of claim 1, wherein the *hedgehog* antagonist is administered as part of a therapeutic or cosmetic application.

11. The method of claim 10, wherein the therapeutic or cosmetic application is selected from the group consisting of regulation of neural tissues, bone and cartilage formation and repair, regulation of spermatogenesis, regulation of smooth muscle, regulation of lung, liver and other organs arising from the primative gut, regulation of hematopoietic function, regulation of skin and hair growth, etc.

12. A method for inhibiting an altered growth state of a cell having a *hedgehog* gain-of-function phenotype, comprising contacting the cell with a *hedgehog* antagonist in a sufficient amount to inhibit the altered growth state, wherein the *hedgehog* antagonist is a organic molecule represented in the general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

X is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-(R_8)N-N(R_8)$-, $-ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y is selected from -C(=O)-, -C(=S)-, $-S(O_2)$-, -S(O)-, -C(=NCN)-, $-P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-R_8N-NR_8$-, $-ONR_8$-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

13. The method of claim 12, wherein R_1 represents a substituted or unsubstituted aryl or heteroaryl group.

14. The method of claim 12, wherein X-Y-Z taken together represents a urea or an amide.

15. The method of claim 12, wherein W is a substituted or unsubstituted benzene ring.

16. The method of claim 12, wherein X represents a diazacarbocycle.

17. The method of claim 12, wherein R_2 represents a substituted or unsubstituted aryl or heteroaryl group.

18. The method of claim 12, wherein R_8, independently for each occurrence, is selected from H and lower alkyl.

19. The method of claim 12, wherein X is selected from -N(R_8)-, -O-, -S-, and a direct bond; Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-; and Z is selected from -N(R_8)-, -O-, -S-, and a direct bond, such that at least one of X and Z is present.

20. The method of claim 12, wherein at least one of X and Z is present.

21. The method of claim 12, wherein Y is selected from -C(=O)-, -C(=S)-, and -S(O_2)-.

22. The method of claim 12, wherein the *hedgehog* antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 mM or less.

23. The method of claim 12, wherein the hedgehog antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 μM or less.

24. The method of claim 12, wherein the hedgehog antagonist inhibits *hedgehog*-mediated signal transduction with an ED_{50} of 1 nM or less.

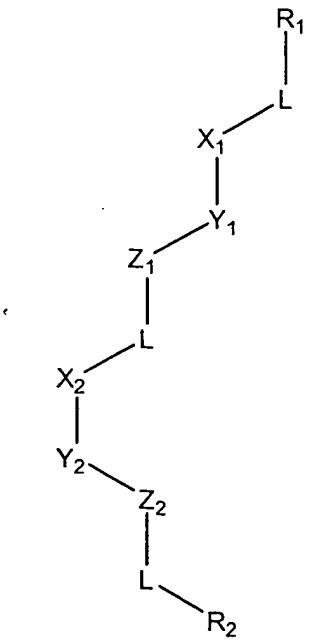
25. The method of claim 12, wherein the cell is contacted with the *hedgehog* antagonist *in vitro*.

26. The method of claim 12, wherein the cell is contacted with the *hedgehog* antagonist *in vivo*.

27. The method of claim 12, wherein the *hedgehog* antagonist is administered as part of a therapeutic or cosmetic application.

28. The method of claim 27, wherein the therapeutic or cosmetic application is selected from the group consisting of regulation of neural tissues, bone and cartilage formation and repair, regulation of spermatogenesis, regulation of smooth muscle, regulation of lung, liver and other organs arising from the primative gut, regulation of hematopoietic function, regulation of skin and hair growth, etc.

29. A pharmaceutical preparation comprising a sterile pharmaceutical excipient and a compound represented by the general formula (I):



Formula I

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X_1 and X_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -(R_8)N-N(R_8)-, -ON(R_8)-, a heterocycle, or a direct bond between L and Y_1 or Y_2, respectively;

Y_1 and Y_2 are selected, independently, from -C(=O)-, -C(=S)-, -$S(O_2)$-, -S(O)-, -C(=NCN)-, -P(=O)(OR_2)-, a heteroaromatic group, or a direct bond between X_1 and Z_1 or X_2 and Z_2, respectively;

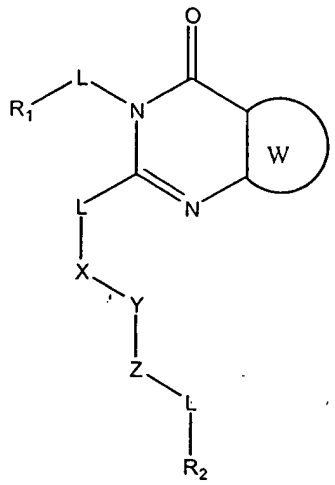
Z_1 and Z_2 are selected, independently, from -$N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y_1 or Y_2, respectively, and L;

R_8, independently for each occurrence, represents H, lower alkyl, -$(CH_2)_n$aryl (substituted or unsubstituted), -$(CH_2)_n$heteroaryl (substituted or unsubstituted), or two R_8 taken together form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

30. A pharmaceutical preparation comprising a sterile pharmaceutical excipient and a compound represented by the general formula (II):



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H; lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted);

L, independently for each occurrence, is absent or represents $-(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, $-(CH_2)_n$alkenyl-, $-(CH_2)_n$alkynyl-, $-(CH_2)_nO(CH_2)_p$-, $-(CH_2)_nNR_2(CH_2)_p$-, $-(CH_2)_nS(CH_2)_p$-, $-(CH_2)_n$alkenyl$(CH_2)_p$-, $-(CH_2)_n$alkynyl$(CH_2)_p$-, $-O(CH_2)_n$-, $-NR_2(CH_2)_n$-, or $-S(CH_2)_n$-;

X is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-(R_8)N-N(R_8)$-, $-ON(R_8)$-, a heterocycle, or a direct bond between L and Y;

Y is selected from -C(=O)-, -C(=S)-, $-S(O_2)$-, -S(O)-, -C(=NCN)-, $-P(=O)(OR_2)$-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from $-N(R_8)$-, -O-, -S-, -Se-, -N=N-, -ON=CH-, $-R_8N-NR_8$-, $-ONR_8$-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, aryl (substituted or unsubstituted), aralkyl (substituted or unsubstituted), heteroaryl (substituted or unsubstituted), or heteroaralkyl (substituted or unsubstituted), or two R_8 taken together

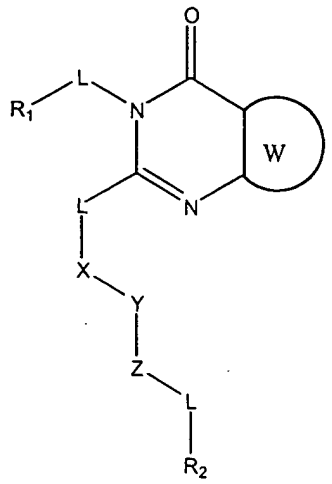
form a 4- to 8-membered ring, together with the atoms to which they are attached, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10; and

n, individually for each occurrence, represents an integer from 0 to 10.

31. A compound having the general structure of Formula II:



Formula II

wherein, as valence and stability permit,

R_1 and R_2, independently for each occurrence, represent H, lower alkyl, -$(CH_2)_n$aryl (e.g., substituted or unsubstituted), or -$(CH_2)_n$heteroaryl (e.g., substituted or unsubstituted);

L, independently for each occurrence, is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

X is -NH-;

Y is selected from -C(=O)-, -C(=S)-, -S(O_2)-, -S(O)-, -C(=NCN)-, -P(=O)(OR_2)-, a heteroaromatic group, or a direct bond between X and Z;

Z is selected from -N(R_8)-, -O-, -S-, -Se-, -N=N-, -ON=CH-, -R_8N-NR_8-, -ONR_8-, a heterocycle, or a direct bond between Y and L;

R_8, independently for each occurrence, represents H, lower alkyl, -(CH_2)$_n$aryl (e.g., substituted or unsubstituted), -(CH_2)$_n$heteroaryl (e.g., substituted or unsubstituted), or two R_8 taken together may form a 4- to 8-membered ring, e.g., with X_1 and Z_1 or X_2 and Z_2, which ring may include one or more carbonyls;

W represents a substituted or unsubsituted aryl or heteroaryl ring fused to the pyrimidone ring;

p represents, independently for each occurrence, an integer from 0 to 10, preferably from 0 to 3; and

n, individually for each occurence, represents an integer from 0 to 10, preferably from 0 to 5.

32. The compound of claim 31, wherein L adjacent to X represents -(unbranched lower alkyl)-.

33. The compound of claim 31, wherein R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, aryl, hydroxyl, (unbranched alkyl-O-), silyloxy, amino, nitro, thiol, amino, imino, amido, phosphoryl, phosphonate, phosphine, carbonyl, carboxyl, carboxamide, anhydride, silyl, thioether, alkylsulfonyl, arylsulfonyl, sulfoxide, selenoether, ketone, aldehyde, ester, or -(CH_2)$_m$-R_8.

34. The compound of claim 31, wherein R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, aryl, nitro, thiol, imino, amido, carbonyl, carboxyl, anhydride, thioether, alkylsulfonyl, arylsulfonyl, ketone, aldehyde, and ester.

35. The compound of claim 31, wherein R_1 represents an unsubstituted aryl or heteroaryl ring, or an aryl or heteroaryl ring substituted with substituents selected from H, halogen, cyano, alkyl, alkenyl, alkynyl, nitro, amido, carboxyl, anhydride, alkylsulfonyl, ketone, aldehyde, and ester.

36. The compound of any of claims 31-35, wherein R_2 is a substituted or unsubstituted aryl or heteroaryl ring.

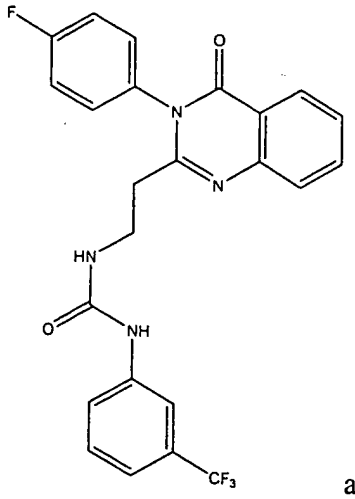
37. The compound of any of claims 31-35, wherein X-Y-Z- represents an amide or urea linkage.

38. The compound of 38, wherein R_8 represents H for all occurrences.

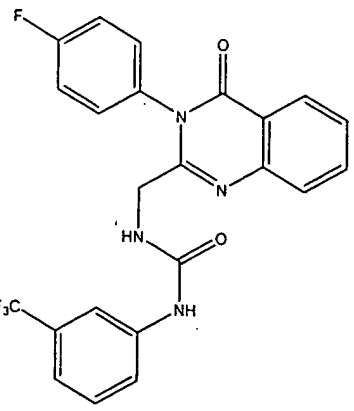
39. The compound of claim 37, wherein R_2 is a substituted or unsubstituted aryl or heteroaryl ring.

40. The compound of claim 37, wherein L is absent adjacent to R_1.

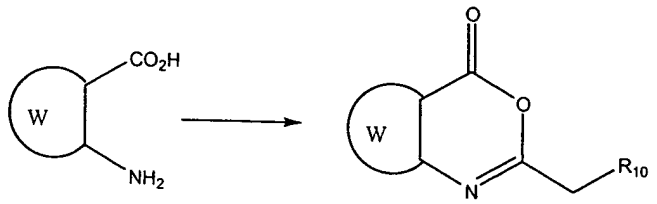
41. A compound having a structure selected from the structures depicted in Figures 32j, k, l,



and



42. A method for preparing a bicyclic compound, comprising heating a reaction mixture comprising a compound having a structure of Formula X with a carboxylic acid anhydride according to the scheme:



X XI

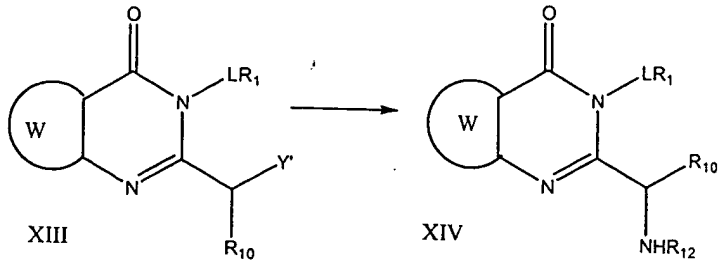
wherein W represents an aryl or heteroaryl ring, and

R_{10} represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl, and

wherein the anhydride has a structure of $(R_{10}CH_2C{=}O)_2O$, wherein R' is CH_2R_{10}.

43. The method of claim 42, wherein the reaction mixture consists essentially of the compound of Formula X and the carboxylic acid anhydride at the start of the reaction.

44. A method for preparing an amine, comprising contacting a compound having a structure of Formula XIII with an amine having a structure of H_2NR_{12} according to the scheme:



wherein W represents an aryl or heteroaryl ring;

R_1 represents H, lower alkyl, -$(CH_2)_n$aryl, or -$(CH_2)_n$heteroaryl;

L is absent or represents -$(CH_2)_n$-alkyl, -alkenyl-, -alkynyl-, -$(CH_2)_n$alkenyl-, -$(CH_2)_n$alkynyl-, -$(CH_2)_nO(CH_2)_p$-, -$(CH_2)_nNR_2(CH_2)_p$-, -$(CH_2)_nS(CH_2)_p$-, -$(CH_2)_n$alkenyl$(CH_2)_p$-, -$(CH_2)_n$alkynyl$(CH_2)_p$-, -$O(CH_2)_n$-, -$NR_2(CH_2)_n$-, or -$S(CH_2)_n$-;

Y' represents a halogen;

R_{10} represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl; and

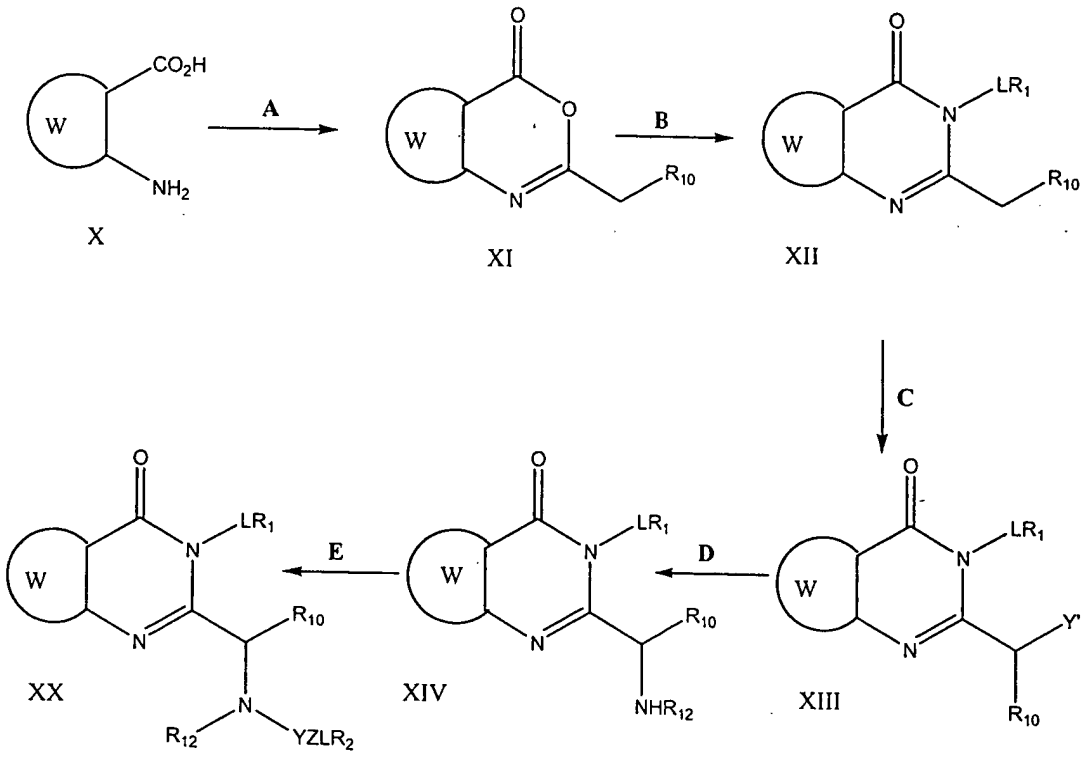
R_{12} represents a lower alkyl group or a silyl group,

wherein the amine and the compound having a structure of Formula XIII are combined with a polar solvent comprising less than about 50% water.

45. The method of claim 44, wherein the polar solvent comprises an alcohol.

46. The method of claim 45, wherein the alcohol is selected from methanol, ethanol, propanol, isopropanol, butanol, isobutanol, t-butanol, sec-butanol, ethylene glycol, and 1,3-propanediol.

47. A method for preparing a compound, comprising performing reactions according to the scheme:



wherein step (A) comprises reacting a compound having a structure of Formula X, wherein W represents a substituted or unsubstituted aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid group in adjacent (*ortho*) positions, with an acylating agent having the formula $R_{10}CH_2C(=O)X'$, wherein R_{10}, independently for each occurrence, represents substituted or unsubstituted alkyl, alkenyl, cycloalkyl, cycloalkylalkyl, aralkyl, or heteroaralkyl, and X' represents a halogen or $-OC(=O)CH_2R_{10}$, under conditions that produce a compound having a structure of Formula XI;

step (B) comprises reacting a compound having a structure of Formula XI with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XII;

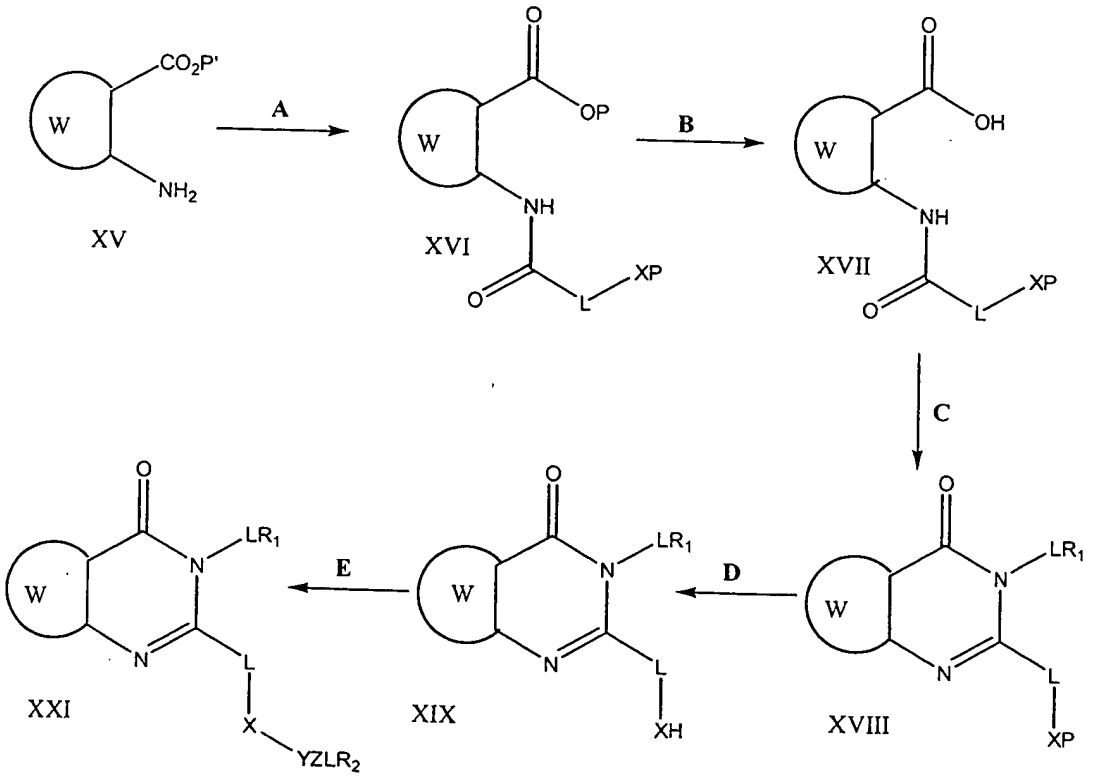
step (C) comprises reacting a compound having a structure of Formula XII with a halogenating agent, such as chlorine, bromine, iodine, N-bromosuccinimide, N-

chlorosuccinimide, N-iodosuccinimide, ClBr, IBr, ClI, or a reagent that generates a halogen radical (such as Cl· , Br· , or I·) under conditions that result in a compound having a structure of Formula XIII, wherein Y' represents a halogen such as Cl, Br, or I;

step (D) comprises reacting a compound having a structure of Formula XIII with an amine having the formula H_2NR_{12}, wherein R_{12} represents a lower alkyl group or a silyl group, such as a trialkylsilyl, triarylsilyl, dialkylarylsilyl, or diarylalkylsilyl group, under conditions that result in a compound having a structure of Formula XIV; and

step (E) comprises reacting a compound having a structure of Formula XIV with a terminating group having a structure of R_2V' to produce a compound having a structure of Formula XX, wherein R_2 is as defined above, and V' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, $ZC(=W)WC(=W)ZR_2$, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an activated acylating moiety prepared *in situ*.

48. A method for preparing a compound, comprising performing reactions according to the scheme:



wherein step (A) comprises reacting a compound having a structure of Formula XV, wherein P' represents H or a protecting group, W represents a substituted or unsubstituted aryl or heteroaryl ring, such as a benzene ring, having an amino group and a carboxylic acid or ester group in adjacent (*ortho*) positions, with an acylating agent having the formula PXLC(=O)X', wherein X and L are as defined above, P represents a protecting group, and X' represents a halogen, –OC(=O)LXP, or a functional group generated by reacting a carboxyl group with an activating agent, such as a carbodiimide (e.g., diisopropylcarbodiimide, dicyclohexylcarbodiimide, 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide, etc.), phosphorous-based reagents (such as BOP-Cl, PyBROP, etc.), oxalyl chloride, phosgene, triphosgene, carbonyldiimidazole, or any other reagent that reacts with a carboxylic acid group resulting in a reactive intermediate having an increased susceptibility, relative to the carboxylic acid, towards coupling with an amine, under conditions that produce a compound having a structure of Formula XVI;

step (B) comprises deprotecting the ester of a compound having a structure of Formula XVI to produce a carboxylic acid having a structure of Formula XVII, if necessary;

step (C) comprises reacting a compound having a structure of Formula XVII with an amine having the formula R_1LNH_2, wherein R_1 and L are as defined above, under conditions that result in a compound having a structure of Formula XVIII;

step (D) comprises removing the protecting group P from a compound having a structure of Formula XVIII to generate a compound having a structure of Formula XIX;

step (E) comprises reacting a compound having a structure of Formula XIX with a terminating group having a structure of R_2Y' to produce a compound having a structure of Formula XXI, wherein R_2 is as defined above, and Y' represents a functional group selected from ZC(=W)Cl, ZC(=W)Br, isocyanate, isothiocyanate, $ZC(=W)WC(=W)ZR_2$, ZSO_2Cl, ZSO_2Br, ZSOCl, ZSOBr, or an active acylating moiety prepared *in situ*.

Abstract

The present invention makes available methods and reagents for inhibiting aberrant growth states resulting from *hedgehog* gain-of-function by contacting the cell with a *hedgehog* antagonist, such as a small molecule, in a sufficient amount to aberrant growth state, e.g., to agonize a normal *ptc* pathway or antagonize *hedgehog* activity.